

# GLOBAL REGULARITY OF WAVE MAPS V. LARGE DATA LOCAL WELLPOSEDNESS AND PERTURBATION THEORY IN THE ENERGY CLASS

TERENCE TAO

ABSTRACT. Using the harmonic map heat flow and the function spaces of Tataru [28] and the author [20], we establish a large data local well-posedness result in the energy class for wave maps from two-dimensional Minkowski space  $\mathbf{R}^{1+2}$  to hyperbolic spaces  $\mathbf{H}^m$ . This is one of the five claims required in [24] to prove global regularity for such wave maps.

## 1. INTRODUCTION

**1.1. The energy space.** This paper is a technical component of a larger program [24] to establish large data global regularity for the initial value problem for two-dimensional wave maps into hyperbolic spaces. Specifically, we establish in this paper a large data local well-posedness result for this problem in the energy space  $\dot{\mathcal{H}}^1$  constructed in [25], which is a key ingredient in the arguments of [24].

We begin by recalling some key features of this energy space. Fix  $m \geq 1$ ; we allow all implied constants to depend on  $m$ . Let  $\mathbf{H} = (\mathbf{H}^m, h) \equiv SO(m, 1)/SO(m)$  be the  $m$ -dimensional *hyperbolic space*, i.e. the simply-connected  $m$ -dimensional Riemannian manifold of constant negative sectional curvature  $-1$ . Here  $SO(m, 1)$  and  $SO(m)$  are the special orthogonal groups of Minkowski space  $\mathbf{R}^{1+m}$  and Euclidean space  $\mathbf{R}^m$  respectively. We define *classical data* to be a pair  $\Phi = (\phi_0, \phi_1)$ , where  $\phi_0 : \mathbf{R}^2 \rightarrow \mathbf{H}$  is a smooth map which differs from a constant  $\phi_0(\infty)$  by a Schwartz function (where we embed  $\mathbf{H}$  in  $\mathbf{R}^{1+m}$  to define the Schwartz space), and  $\phi_1 : \mathbf{R}^2 \rightarrow T\mathbf{H}$  is a Schwartz function such that  $\phi_1(x)$  lies in the tangent plane  $T_{\phi_0(x)}\mathbf{H}$  of  $\mathbf{H}$  at  $\phi_0(x)$  for every  $x \in \mathbf{R}^2$ , and let  $\mathcal{S}$  be the space of all classical data; this can be given the structure of a topological space by using the Schwartz topology. With regards to wave maps, one should interpret  $\phi_0$  and  $\phi_1$  as being the initial position and initial velocity respectively of a (classical) wave

map at some time. We observe the Lorentz rotation symmetry

$$\text{Rot}_U : (\phi_0(x), \phi_1(x)) \mapsto (U\phi_0(x), dU(\phi_0(x))(\phi_1(x))) \quad (1)$$

for any Lorentz rotation  $U \in SO(m, 1)$  that acts continuously on  $\mathcal{S}$ .

Given any classical initial data  $\Phi = (\phi_0, \phi_1)$ , one can define the *energy*

$$E(\Phi) := \frac{1}{2} \int_{\mathbf{R}^2} |\partial_x \phi_0|_{\phi_0^* h}^2 + |\phi_1|_{\phi_0^* h}^2 dx \quad (2)$$

where  $\phi_0^* h$  is the pullback of the metric  $h$  by  $\phi_0$ . The energy functional  $E$  is a continuous map from  $\mathcal{S}$  to  $[0, +\infty)$ .

In [25], an energy space  $\dot{\mathcal{H}}^1$  was constructed with the following properties<sup>1</sup>:

**Theorem 1.2** (Energy space). [25] *There exists a complete metric space  $\dot{\mathcal{H}}^1$  with a continuous map  $\iota : \mathcal{S} \rightarrow \dot{\mathcal{H}}^1$ , that obeys the following properties:*

- (i)  $\iota(\mathcal{S})$  is dense in  $\dot{\mathcal{H}}^1$ .
- (ii)  $\iota$  is invariant under the action (1) of the Lorentz rotation group  $SO(m, 1)$ , thus  $\iota(\text{Rot}_U \Phi) = \iota(\Phi)$  for all  $\Phi \in \mathcal{S}$ . Conversely, if  $\iota(\Phi) = \iota(\Psi)$ , then  $\Psi = \text{Rot}_U(\Phi)$  for some  $U \in SO(m, 1)$ .
- (iii) The energy functional  $E : \mathcal{S} \rightarrow [0, +\infty)$  extends to a continuous map  $E : \dot{\mathcal{H}}^1 \rightarrow [0, +\infty)$  (after quotienting out by Lorentz rotations).
- (iv) For any  $\Phi \in \dot{\mathcal{H}}^1$ , we have  $E(\Phi) = d(\Phi, \text{const})^2$ , where  $\text{const} := \iota(p, 0)$  is the image of the constant map  $(p, 0)$  for any  $p \in \mathbf{H}$ .

*Remark 1.3.* From (i), (ii) one can interpret  $\dot{\mathcal{H}}^1$  as a metric completion of the quotient  $SO(m, 1) \backslash \mathcal{S}$  of  $\mathcal{S}$  by the action of the Lorentz group. The construction of this space  $\dot{\mathcal{H}}^1$  involves the harmonic map heat flow and will be reviewed in Section 5, as the details of this construction will be important in our arguments.

**1.4. Wave maps.** Define a *classical wave map* to be a pair  $(\phi, I)$ , where  $I$  is a time interval, and  $\phi : I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  is a smooth map which differs from a constant  $\phi(\infty) \in \mathbf{H}$  by a Schwartz function in space<sup>2</sup>, and which obeys the equation

$$(\phi^* \nabla)^\alpha \partial_\alpha \phi = 0, \quad (3)$$

<sup>1</sup>In [25], several additional properties of this space were also established, but they will not be needed in this paper and so we have omitted them.

<sup>2</sup>We say that a function  $\phi : I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  is *Schwartz* if it is smooth, and  $\partial_t^j \partial_x^k \phi$  is rapidly decreasing in space for all  $j, k \geq 0$ , uniformly in time. We let  $\text{Schwartz}(I)$  denote the space of all Schwartz functions, with the usual Frechet space topology.

where  $\phi^*\nabla$  is the pullback of the Levi-Civita connection  $\nabla$  on  $\mathbf{H}$  by  $\phi$ . Here and in the sequel we use Greek indices  $\alpha, \beta = 0, 1, 2$  to range over the coordinates of Minkowski space  $\mathbf{R}^{1+2}$  (with the usual metric  $-dt^2 + dx_1^2 + dx_2^2$ ), with the usual summation, raising, and lowering conventions (and setting  $x_0 := t$ ). Roman indices  $i, j = 1, 2$  will be summed over the spatial coordinates only (with no attempt to raise or lower coordinates).

Observe that for any time  $t \in I$ , the data  $\phi[t] := (\phi(t), \partial_t \phi(t))$  of a classical wave map  $(\phi, I)$  lies in  $\mathcal{S}$ , and indeed  $\phi$  can be viewed as a smooth curve  $\phi : I \rightarrow \mathcal{S}$ . We refer to [5], [8], [15], [17], [30], [22, Chapter 6], [13], [11] for surveys of the initial value problem for wave maps, which is of course the primary concern of this project. As is well known, wave maps have a conserved energy  $E(\phi) = E(\phi(t))$ , which in this two spatial dimensional setting is also invariant under the natural scale invariance  $\phi^{(\lambda)}(t, x) := \phi(\frac{t}{\lambda}, \frac{x}{\lambda})$  for this problem.

The purpose of this paper is to establish the following large data local well-posedness result for wave maps in the energy class:

**Theorem 1.5** (Large data local-wellposedness in the energy space). *For every time  $t_0 \in \mathbf{R}$  and every initial data  $\Phi_0 \in \mathcal{H}^1$  there exists a maximal lifespan  $I \subset \mathbf{R}$ , and a maximal Cauchy development  $\phi : t \mapsto \phi[t]$  from  $I \rightarrow \mathcal{H}^1$ , which obeys the following properties:*

- (i) (Local existence)  $I$  is an open interval<sup>3</sup> containing  $t_0$ .
- (ii) (Strong solution)  $\phi : I \rightarrow \mathcal{H}^1$  is continuous.
- (iii) (Persistence of regularity) If  $\Phi_0 = \iota(\tilde{\Phi}_0)$  for some classical data  $\tilde{\Phi}_0$ , then there exists a classical wave map  $(\tilde{\phi}, I)$  with initial data  $\tilde{\phi}[t_0] = \tilde{\Phi}_0$  such that  $\phi[t] = \iota(\tilde{\phi}[t])$  for all  $t \in I$ .
- (iv) (Continuous dependence) If  $\Phi_{0,n}$  is a sequence of data in  $\mathcal{H}^1$  converging to a limit  $\Phi_{0,\infty}$ , and  $\phi_n : I_n \rightarrow \mathcal{H}^1$  and  $\phi_\infty : I_\infty \rightarrow \mathcal{H}^1$  are the associated maximal Cauchy developments on the associated maximal lifespans, then for every compact subinterval  $K$  of  $I_\infty$ , we have  $K \subset I_n$  for all sufficiently large  $n$ , and  $\phi_n$  converges uniformly to  $\phi$  on  $K$  in the  $\mathcal{H}^1$  topology.
- (v) (Maximality) If  $t_* \in \mathbf{R}$  is a finite endpoint of  $I$ , then  $\phi(t)$  has no convergent subsequence in  $\mathcal{H}^1$  as  $t \rightarrow t_*$ .

---

<sup>3</sup>An *interval* in this paper is a connected subset of  $\mathbf{R}$  with non-empty interior. It can be closed, open, or half-open, and it can be bounded or unbounded.

This result is one of five claims required in [24] to establish global regularity for classical wave maps; of the other four claims, three<sup>4</sup> of them (including the construction of the energy space, Theorem 1.2) were proven in [25]; the remaining claim, concerning the existence of almost periodic maximal Cauchy developments in the case when global regularity fails, will be established in the sequels [26], [27] to this paper.

Theorem 1.5 has the following corollary:

**Corollary 1.6** (Small data global-wellposedness in the energy space). *There exists an absolute constant  $\varepsilon_0 > 0$  (depending only on  $m$ ) for which the following statements hold: given every time  $t_0 \in \mathbf{R}$  and every initial data  $\Phi_0 \in \dot{\mathcal{H}}^1$  with  $E(\Phi_0) < \varepsilon_0$  there exists a global Cauchy development  $\phi[t]$  from  $\mathbf{R}$  to  $\dot{\mathcal{H}}^1$ , which obeys the following properties:*

- (i) (Strong solution)  $\phi : \mathbf{R} \rightarrow \dot{\mathcal{H}}^1$  is continuous.
- (ii) (Persistence of regularity) If  $\Phi_0 = \iota(\tilde{\Phi}_0)$  for some classical data  $\tilde{\Phi}_0$ , then there exists a classical wave map  $(\tilde{\phi}, \mathbf{R})$  with initial data  $\tilde{\phi}[t_0] = \tilde{\Phi}_0$  such that  $\phi[t] = \iota(\tilde{\phi}[t])$  for all  $t \in \mathbf{R}$ .
- (iii) (Continuous dependence) If  $\Phi_{0,n}$  is a sequence of data in  $\dot{\mathcal{H}}^1$  with  $E(\Phi_{0,n}) < \varepsilon_0$  converging to a limit  $\Phi_{0,\infty}$  with  $E(\Phi_{0,\infty}) < \varepsilon_0$ , and  $\phi_n : I_n \rightarrow \dot{\mathcal{H}}^1$  and  $\phi_\infty : I_\infty \rightarrow \dot{\mathcal{H}}^1$  are the associated global Cauchy developments, then  $\phi_n$  converges locally uniformly to  $\phi$  on compact intervals in the  $\dot{\mathcal{H}}^1$  topology.

*Proof.* In view of Theorem 1.5, the only claim that needs verification is that the maximal lifespan of any data of sufficiently small energy is global. Assume for contradiction that this were not the case, then one can find a sequence  $\Phi_n \in \dot{\mathcal{H}}^1$  with energy going to zero, whose maximal Cauchy developments  $\phi_n : I_n \rightarrow \dot{\mathcal{H}}^1$  from the initial time  $t_n$  were not global. By time translation symmetry<sup>5</sup> we can take  $t_n = 0$ ; by scaling symmetry one can ensure that  $\text{dist}(\partial I_n, 0) \rightarrow 0$  as  $n \rightarrow \infty$ . By Theorem 1.2(iv), we know that  $\Phi_n$  converges to constant data, which of course has a global classical solution to the wave maps problem. But this now contradicts Theorem 1.5(iii).  $\square$

Corollary 1.6, when specialised to classical data, gives a global regularity result for small energy:

---

<sup>4</sup>One of these claims, namely the non-existence of non-trivial travelling or self-similar wave maps, was only proven conditionally on Theorem 1.5. However we will not use that claim in this paper, so the arguments are not circular.

<sup>5</sup>See [24] for a justification of the various symmetries of the wave maps equation in the energy class.

**Corollary 1.7** (Small energy regularity). *Let  $\varepsilon_0$  be as in the preceding corollary. If  $\Phi \in \mathcal{S}$  with  $E(\Phi) < \varepsilon_0$  and  $t_0 \in \mathbf{R}$ , then there is a unique global classical wave map  $\phi : \mathbf{R} \rightarrow \mathcal{S}$  with  $\phi[t_0] = \Phi$ .*

*Remark 1.8.* Corollary 1.7 for the case  $m = 2$  of the hyperbolic plane target was established in [10] (see also [20], [29] for closely related results in other manifolds). It is likely that one can partially reverse the above implications and deduce Theorem 1.5 from Theorem 1.6 by exploiting finite speed of propagation (see [16] for an example of such an argument for the energy-critical wave equation, or Section 3 below for such an argument in the subcritical regularity setting). However there is a technical obstacle to doing so, namely that the energy topology that we use for  $\dot{\mathcal{H}}^1$  is constructed via the harmonic map heat flow, and so it is not immediately obvious that this topology is compatible with localisation in space. While it is likely that this obstacle could be resolved with additional effort, we have chosen not to do so here.

*Remark 1.9.* Very recently, a small energy global well-posedness result for Schrödinger maps into the sphere  $S^2$  was established in [2]. There are many common features between the results here and those in [2], most notably the reliance on the caloric gauge. However, the function spaces used for the Schrödinger map and wave map equation are very different from each other; also, there are additional technical difficulties in setting up the iteration scheme in the large energy case which can be avoided in the small energy setting.

**1.10. Organisation of the paper.** The large data local wellposedness claims in Theorem 1.5 will be proven as follows. After setting out our basic notation in Section 2, we shall establish a subcritical local well-posedness theory in Section 3. In Section 4, we use this theory to reduce matters to the task of establishing some abstract *a priori* estimates for classical wave maps (Theorem 4.1). In order to prove this theorem, we introduce the caloric gauge in Section 5, and reduce matters to establishing analogous *a priori* estimates for classical wave maps in the caloric gauge (Theorem 5.11).

To achieve this, we develop some abstract parabolic regularity theory in Section 6 and use it to obtain good control on the heat flow at the initial time in Section 7. Then, in Section 8 we abstractly describe the hyperbolic function spaces we will need for the *a priori* estimates (Theorem 8.1). In Section 9 we use the estimates in Theorem 8.1 (and the abstract parabolic regularity theory mentioned earlier) to establish Theorem 5.11. Finally, in Section 8 we recall the spaces constructed in [20] which will allow us to verify Theorem 8.1 relatively quickly.

**1.11. Acknowledgements.** This project was started in 2001, while the author was a Clay Prize Fellow. The author thanks Andrew Hassell and the Australian National University for their hospitality when a substantial portion of this work was initially conducted, and to Ben Andrews and Andrew Hassell for a crash course in Riemannian geometry and manifold embedding, and in particular to Ben Andrews for explaining the harmonic map heat flow. The author also thanks Mark Keel for background material on wave maps, Daniel Tataru for sharing some valuable insights on multilinear estimates and function spaces, and to Igor Rodnianski and Jacob Sterbenz for valuable discussions. The author is supported by NSF grant DMS-0649473 and a grant from the Macarthur Foundation.

## 2. NOTATION AND BASIC ESTIMATES

In order to efficiently manage the rather large amount of technical computations, it is convenient to introduce a substantial amount of notation, as well as some basic estimates that we will use throughout the paper.

**2.1. Small exponents.** We will need to fix three small exponents

$$0 < \delta_0 \ll \delta_1 \ll \delta_2 \ll 1$$

which are absolute constants, with each  $\delta_i$  being sufficiently small depending on all higher  $\delta_i$ . The exact choice of constants is not important, but for sake of concreteness one could take  $\delta_i := 10^{-10^{3-i}}$ , for instance. All the implied constants in the asymptotic notation below can depend on these exponents.

*Remark 2.2.* The interpretation of these constants in our argument will be as follows. The largest constant  $\delta_2$  is the exponent that quantifies certain useful exponential decays in frequency-localised linear, bilinear, and trilinear estimates when the ratio of two frequencies becomes favorable. The intermediate constant  $\delta_1$  is used to design the weakly frequency localised space  $S_k$  (and its variant  $S_{\mu,k}$ , adapted to the large data theory) that we will encounter later in this paper. The smallest constant  $\delta_0$  is used to control the fluctuation of the frequency envelopes  $c(s)$  that we will use to control solutions.

It will be useful to have some standard cutoff functions that compare two frequency parameters  $k, k'$ :

**Definition 2.3** (Cutoff functions). Given any integers  $k, k'$ , we define  $\chi_{k \geq k'} = \chi_{k' \leq k} := \min(1, 2^{-(k-k')})$  and  $\chi_{k=k'} := 2^{-|k-k'|}$ .

Thus  $\chi_{k' \leq k}$  is weakly localised to the region  $k' \leq k$ , and similarly for the other cutoffs. In practice we shall usually raise these cutoffs to an exponent such as  $\pm\delta_0$ ,  $\pm\delta_1$ , or  $\pm\delta_2$ .

**2.4. Asymptotic notation.** The dimension  $m$  of the target hyperbolic space  $\mathbf{H}^m$  is fixed throughout the paper, and all implied constants can depend on  $m$ .

We use  $X = O(Y)$  or  $X \lesssim Y$  to denote the estimate  $|X| \leq CY$  for some absolute constant  $C > 0$ , that can depend on the  $\delta_i$  and the dimension  $m$  of the target hyperbolic space. If we wish to permit  $C$  to depend on some further parameters, we shall denote this by subscripts, e.g.  $X = O_k(Y)$  or  $X \lesssim_k Y$  denotes the estimate  $|X| \leq C_k Y$  where  $C_k > 0$  depends on  $k$ .

Note that parameters can be other mathematical objects than numbers. For instance, the statement that a function  $u : \mathbf{R}^2 \rightarrow \mathbf{R}$  is Schwartz is equivalent to the assertion that one has a bound of the form  $|\partial_x^k u(x)| \lesssim_{j,k,u} \langle x \rangle^{-j}$  for all  $j, k \geq 0$  and  $x \in \mathbf{R}^2$ , where  $\langle x \rangle := (1 + |x|^2)^{1/2}$ .

**2.5. Schematic notation.** We use  $\partial_x$  as an abbreviation for  $(\partial_1, \partial_2)$ , thus for instance

$$\begin{aligned} |\partial_x \phi|_{\phi^* h}^2 &= \langle \partial_i \phi, \partial_i \phi \rangle_{\phi^* h}^2 \\ |(\phi^* \nabla)_x \partial_x \phi| &= \langle (\phi^* \nabla)_i \partial_j \phi, (\phi^* \nabla)_i \partial_j \phi \rangle_{\phi^* h}^2 \end{aligned}$$

with the usual summation conventions.

We use juxtaposition to denote tensor product; thus for instance if  $\psi_x := (\psi_1, \psi_2)$ , then  $\psi_x^2 = \psi_x \psi_x$  denotes the rank 2 tensor with the four components  $\psi_i \psi_j$  for  $i, j = 1, 2$ ; similarly,  $\partial_x^2$  is the rank 2 tensor operator with four components  $\partial_i \partial_j$  for  $i, j = 1, 2$ , and so forth.

If  $X$  is a tensor-valued quantity, we use  $\mathcal{O}(X)$  to denote an expression which is *schematically* of the form  $X$ , which means that it is a tensor whose components are linear combinations of those in  $X$ , with coefficients being absolute constants (depending only on  $m$ ). Thus for instance, if  $v, w \in \mathbf{R}^m$ , then the anti-symmetric matrix  $v \wedge w := vw^\dagger - wv^\dagger$  has the schematic form  $v \wedge w = \mathcal{O}(vw)$ . If the coefficients in the schematic representation depend on a parameter, we will denote this by subscripts. Thus for instance we have the *Leibniz rule*

$$\partial_x^j \mathcal{O}(\phi\psi) = \sum_{j_1, j_2 \geq 0: j_1 + j_2 = j} \mathcal{O}_j(\partial_x^{j_1} \phi \partial_x^{j_2} \psi) \quad (4)$$

and similarly for products of three or more functions.

**2.6. Difference notation.** Throughout this paper, we adopt the notational conventions  $\delta f := f' - f$  and  $f^* := (f, f')$  for any field  $f$  for which the right-hand side makes sense. For future reference we observe the discretised Leibniz rule

$$\delta(fg) = \mathcal{O}((\delta f)g^*) + \mathcal{O}(f^*\delta g) \quad (5)$$

and similarly for products of three or more functions.

**2.7. Function spaces.** We use the usual Lebesgue spaces  $L_x^p(\mathbf{R}^2)$  and Sobolev spaces  $H_x^s(\mathbf{R}^2)$ , and create spacetime norms such as  $L_t^q L_x^p(I \times \mathbf{R}^2)$  in the usual manner. Later on we shall also use the more complicated spaces adapted to the wave maps problem from [20] (see also [28]).

If  $X$  is a normed space for scalar-valued functions, we also extend  $X$  to functions  $\phi := (\phi_1, \dots, \phi_m)$  taking values in a standard finite-dimensional vector space such as  $\mathbf{R}^m$  with the convention

$$\|\phi\|_X := \left( \sum_{j=1}^m \|\phi_j\|_X^2 \right)^{1/2}.$$

Note that if  $X$  was already a Hilbert space on scalar functions, it continues to be a Hilbert space on vector-valued functions, and the orthogonal group  $O(m)$  on that space acts isometrically on this space. If  $X$  is merely a normed vector space, then the orthogonal group is no longer isometric, but the action of an element of this group has operator norm bounded above and below by constants depending only on  $m$ .

**Definition 2.8** (Littlewood-Paley projections). Let  $\varphi(\xi)$  be a radial bump function supported in the ball  $\{\xi \in \mathbf{R}^2 : |\xi| \leq \frac{11}{10}\}$  and equal to 1 on the ball  $\{\xi \in \mathbf{R}^2 : |\xi| \leq 1\}$ . For each integer  $k$ , we define the Fourier multipliers

$$\begin{aligned} \widehat{P_{\leq k} f}(\xi) &:= \varphi(\xi/2^k) \hat{f}(\xi) \\ \widehat{P_{> k} f}(\xi) &:= (1 - \varphi(\xi/2^k)) \hat{f}(\xi) \\ \widehat{P_k f}(\xi) &:= (\varphi(\xi/2^k) - \varphi(2\xi/2^k)) \hat{f}(\xi). \end{aligned}$$

We similarly define  $P_{< k}$  and  $P_{\geq k}$ .

**2.9. The linear heat equation.** Throughout the paper we use  $\Delta := \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$  to denote the (spatial) Laplacian on  $\mathbf{R}^2$ . We use  $e^{s\Delta}$  for  $s > 0$  to denote the free heat propagator

$$e^{s\Delta} u(x) := \frac{1}{4\pi s} \int_{\mathbf{R}^2} e^{-|x-y|^2/4s} u(y) \, dy. \quad (6)$$



From Young's inequality we easily establish the parabolic regularity estimate

$$\|\partial_x^j e^{s\Delta} u\|_{L_x^q(\mathbf{R}^2)} \lesssim_{p,q,j} s^{\frac{1}{p}-\frac{1}{q}+\frac{j}{2}} \|e^{s\Delta} u\|_{L_x^p(\mathbf{R}^2)} \quad (7)$$

valid for all  $s > 0$ ,  $j \geq 0$ , and  $1 \leq p \leq q \leq \infty$ .

We recall *Duhamel's formula*

$$u(s_1) = e^{(s_1-s_0)\Delta} u(s_0) + \int_{s_0}^{s_1} e^{(s_1-s)\Delta} (\partial_s u - \Delta u)(s) ds \quad (8)$$

for any continuous map  $s \mapsto u(s)$  from the interval  $[s_0, s_1]$  to the space of tempered distributions on  $\mathbf{R}^2$ , which can be either scalar or vector valued.

We also need a Strichartz-type estimate:

**Lemma 2.10** (Parabolic Strichartz estimate). *Let  $\psi : \mathbf{R}^+ \times \mathbf{R}^2 \rightarrow \mathbf{R}$  be smooth, with all derivatives uniformly Schwartz in space. Then we have*

$$\left( \int_0^\infty \|\psi\|_{L_x^\infty(\mathbf{R}^2)}^2 ds \right)^{1/2} \lesssim \|\psi(0)\|_{L_x^2(\mathbf{R}^2)} + \int_0^\infty \|(\partial_s - \Delta)\psi(s)\|_{L_x^2(\mathbf{R}^2)} dx.$$

*Proof.* By (8) and Minkowski's inequality it suffices to establish this claim for free solutions to the heat equation, thus  $\psi(s) = e^{s\Delta}\psi(0)$ . But this follows from [25, Lemma 2.5].  $\square$

**2.11. Frequency envelopes.** We first recall a useful definition from [19].

**Definition 2.12** (Frequency envelope). [19] Let  $E > 0$ . A *frequency envelope* of energy  $E$  is a map  $c : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  with

$$\int_0^\infty c(s)^2 \frac{ds}{s} = E \quad (9)$$

such that

$$c(s') \leq \max((s'/s)^{\delta_0}, (s/s')^{\delta_0}) c(s) \quad (10)$$

for all  $s, s' > 0$ .

*Remark 2.13.* The estimate (10) is asserting that  $c(s)$  can grow at most as fast as  $s^{\delta_0}$ , and decay at most as rapidly as  $s^{-\delta_0}$ . In [19] a discretised version of this concept was used, with the continuous variable  $s$  being replaced by the discrete variable  $2^{-2k}$  for integer  $k$ . However, as the heat flow uses a continuous time variable  $s$ , it is more natural to use the continuous version of a frequency envelope. In [29] it was observed that one could take asymmetric envelopes, in which  $c(s)$  is allowed to decay faster as  $s$  decreases than when  $s$  increases. However, due to our use of heat flow (which does not have as good frequency damping

properties near the frequency origin as Littlewood-Paley operators) it is not convenient to use these asymmetric envelopes in our arguments.

*Remark 2.14.* Observe from (9), (10) that if  $c$  is a frequency envelope of energy  $E$ , then we have the pointwise bounds

$$c(s) \lesssim \sqrt{E} \quad (11)$$

for all  $s > 0$ .

We also define the frequency  $k(s) \in \mathbf{Z}$  of a heat-temporal variable  $s > 0$  by the formula

$$k(s) := \lfloor \log_2 s^{-1/2} \rfloor, \quad (12)$$

thus  $2^{-2k(s)} \sim s$ .

We record two useful Gronwall-type inequalities relating to these frequency envelopes. The first lemma involves integral inequalities coming from  $s = +\infty$ :

**Lemma 2.15** (Gronwall-type inequality from  $s = +\infty$ ). *Let  $c$  be a frequency envelope of energy at most  $E$ . Let  $f, g : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  be locally integrable functions with  $\lim_{s \rightarrow \infty} f(s) = 0$  such that  $g$  obeys (10), and*

$$f(s) \leq g(s) + \int_s^\infty f(s')c(s')((s' - s)/s')^{-\theta}(s'/s)^{-3\delta_0} \frac{ds'}{s'}$$

for all  $s > 0$  and some  $0 \leq \theta < 1$ . Then we have

$$f(s) \lesssim_{E,\theta} g(s)$$

for all  $s > 0$ .

*Proof.* We allow implied constants to depend on  $E, \theta$ . If we let  $\tilde{f}(s) := \sup_{s \leq s' \leq 2s} f(s')$ , then we easily compute (using (10)) that

$$\int_s^\infty f(s')c(s')((s' - s)/s')^{-\theta}(s'/s)^{-3\delta_0} \frac{ds'}{s'} \lesssim \int_s^\infty \tilde{f}(s')c(s')(s'/s)^{-3\delta_0} \frac{ds'}{s'}$$

and thus

$$\tilde{f}(s) \lesssim g(s) + \int_s^\infty \tilde{f}(s')c(s')(s'/s)^{-3\delta_0} \frac{ds'}{s'}.$$

Because of this, we see that the  $\theta \neq 0$  case follows from the  $\theta = 0$  case. Henceforth we take  $\theta = 0$ .

Since  $f(s')c(s')$  goes to zero as  $s' \rightarrow \infty$ , we see for  $S$  sufficiently large that

$$f(s) \leq s^{3\delta_0} S^{-3\delta_0} + g(s) + \int_s^S f(s')c(s')(s'/s)^{-3\delta_0} \frac{ds'}{s'}$$

for all  $0 < s \leq S$ . If we set  $G(s) := S^{-3\delta_0} + g(s)s^{-3\delta_0}$  and  $F(s) := \max(f(s)s^{-3\delta_0} - G(s), 0)$ , we thus have

$$F(s) \leq \int_s^S (F(s') + G(s'))c(s') \frac{ds'}{s'}.$$

Applying Gronwall's inequality, we conclude that

$$F(s) \leq \exp\left(\int_s^S c(s') \frac{ds'}{s'}\right) F(S) + \int_s^S \exp\left(\int_s^{s'} c(s'') \frac{ds''}{s''}\right) G(s')c(s') \frac{ds'}{s'}.$$

From (9) and Cauchy-Schwarz we have

$$\int_s^{s'} c(s'') \frac{ds''}{s''} \lesssim \sqrt{\log \frac{s'}{s}}$$

and hence

$$\exp\left(\int_s^{s'} c(s'') \frac{ds''}{s''}\right) \lesssim (s'/s)^{\delta_0}.$$

We thus have

$$F(s) \lesssim (S/s)^{\delta_0} F(S) + \int_s^S (s'/s)^{\delta_0} G(s')c(s') \frac{ds'}{s'}.$$

If we let  $S \rightarrow \infty$ , then  $(S/s)^{\delta_0} F(S) \rightarrow 0$  by hypothesis on  $f$ , and thus

$$f(s)s^{-3\delta_0} - g(s)s^{-3\delta_0} \lesssim \int_s^\infty (s'/s)^{\delta_0} g(s')(s')^{-3\delta_0} c(s') \frac{ds'}{s'}.$$

Using (11) we obtain the claim.  $\square$

By making the change of variables  $s \mapsto 1/s$  we also have a variant from  $s = 0$ :

**Lemma 2.16** (Gronwall-type inequality from  $s = 0$ ). *Let  $c$  be a frequency envelope of energy at most  $E$ . Let  $f, g : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  be locally integrable functions with  $\lim_{s \rightarrow 0} f(s) = 0$  such that  $g$  obeys (10), and*

$$f(s) \leq g(s) + \int_s^\infty f(s')c(s')((s-s')/s)^{-\theta}(s'/s)^{3\delta_0} \frac{ds'}{s'}$$

for all  $s > 0$  and some  $0 \leq \theta < 1$ . Then we have

$$f(s) \lesssim_{E,\theta} g(s)$$

for all  $s > 0$ .

## 3. SUBCRITICAL LOCAL EXISTENCE THEORY

Before we study wave maps in the energy class, we first review the (much simpler) subcritical local existence theory for classical wave maps, in which the bounds are allowed to depend on smoother norms than the energy norm, and in particular on norms which are subcritical with respect to scaling, of  $H^{1+\varepsilon}$  type. This theory was essentially worked out by Klainerman and Selberg [7] (see also the critical Besov space refinement in [28]); the only (slight) wrinkle we need to deal with here is the Lorentz rotation-invariance of the problem, which can distort the local coordinates slightly. This theory is not used in the “quantitative” components of the argument, but is needed for more “qualitative” components, such as ensuring that various continuity arguments can be justified, or that an *a priori* hypothesis can be removed.

For the purposes of this qualitative analysis it is convenient to represent the hyperbolic space  $\mathbf{H} = (\mathbf{H}^m, h)$  concretely in an ambient Minkowski space  $\mathbf{R}^{1+m}$  as the upper unit hyperboloid

$$\mathbf{H} := \{(t, x) \in \mathbf{R}^{1+m} : t = +\sqrt{1 + |x|^2}\} \subset \mathbf{R}^{1+m} \quad (13)$$

with the metric  $dh^2$  induced from the Minkowski metric  $dg^2 = -dt^2 + dx^2$  on  $\mathbf{R}^{1+m}$ , and with the obvious action of  $SO(m, 1)$ . The connection  $\phi^*\nabla$  on the pullback tangent bundle  $\phi^*T\mathbf{H}$  can then be written in these coordinates as

$$(\phi^*\nabla)_i\psi(x) = \partial_i\psi(x) - \langle\psi(x), \partial_i\phi(x)\rangle_{\mathbf{R}^{1+m}}\phi(x). \quad (14)$$

whenever  $\psi$  is Minkowski-orthogonal to  $\phi$  (i.e.  $\langle\psi, \phi\rangle_{\mathbf{R}^{1+m}} = 0$ ), while the wave maps equation (3) in these coordinates becomes

$$\square\phi = \langle\partial^\alpha\phi, \partial_\alpha\phi\rangle_{\mathbf{R}^{1+m}}\phi \quad (15)$$

where  $\square := \partial^\alpha\partial_\alpha$  is the d’Lambertian operator.

We observe the zero tension property

$$(\phi^*\nabla)_i\partial_j\phi = (\phi^*\nabla)_j\partial_i\phi \quad (16)$$

and the constant negative curvature property

$$(\phi^*\nabla)_i(\phi^*\nabla)_j\psi - (\phi^*\nabla)_j(\phi^*\nabla)_i\psi = -(\partial_i\phi \wedge \partial_j\phi)\psi \quad (17)$$

for any section  $\psi$  of  $\phi^*T\mathbf{H}$ , where  $\partial_i\phi \wedge \partial_j\phi \in \Gamma(\text{Hom}(\phi^*T\mathbf{H} \rightarrow \phi^*T\mathbf{H}))$  is the anti-symmetric rank (1, 1) tensor defined by the formula

$$(\partial_i\phi \wedge \partial_j\phi)\psi = \partial_i\phi\langle\partial_j\phi, \psi\rangle_{\phi^*h} - (\partial_j\phi \wedge \partial_i\phi)\psi.$$

We fix a smooth scalar cutoff function  $\eta \in C_0^\infty(\mathbf{R}^2)$  that equals 1 on the ball  $B(0, 1)$  and vanishes outside of  $B(0, 2)$ . Given classical initial

data  $\Phi := (\phi_0, \phi_1) \in \mathcal{S}$  and  $s \geq 1$ , we define the Sobolev-type norms

$$\|\Phi\|_{\mathcal{H}_{\text{loc}}^s} := \|f\|_{L_x^\infty(\mathbf{R}^2)} \quad (18)$$

and

$$\|\Phi\|_{\mathcal{H}^s} := \|\Phi\|_{\mathcal{H}_{\text{loc}}^s} + \|f\|_{L_x^2(\mathbf{R}^2)} \quad (19)$$

where  $f(x_0)$  is the local Sobolev norm near  $x_0$ , defined as

$$f(x_0) := \inf_{U \in SO(m,1)} \|\eta(\cdot - x_0)U(\phi_0)\|_{H_x^s(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U(\phi_1)\|_{H_x^{s-1}(\mathbf{R}^2)}.$$

*Remark 3.1.* These (rather artificial) quantities are not exactly norms (after all,  $\mathcal{S}$  is not a vector space) but should be viewed as nonlinear analogues of the more familiar Sobolev class  $H_x^s(\mathbf{R}^2) \times H_x^{s-1}(\mathbf{R}^2)$  (and their localised counterpart) in the theory of scalar nonlinear wave equations. Observe from construction that these “norms” are invariant with respect to Lorentz rotations.

We have the following (reasonably standard) subcritical local existence result:

**Theorem 3.2** (Subcritical local existence). *If  $\Phi = (\phi_0, \phi_1)$  is classical initial data,  $s > 1$ , and  $t_0 \in \mathbf{R}$ , then there exists a time  $T > 0$  depending only on  $s$  and  $\|\Phi\|_{\mathcal{H}_{\text{loc}}^s}$ , and a classical wave map  $(\phi, [t_0 - T, t_0 + T])$  with initial data  $\phi[t_0] = \Phi$ . Furthermore, we have the persistence of regularity result*

$$\|\phi[t]\|_{\mathcal{H}_{\text{loc}}^{s'}} \lesssim_{s,s',\|\Phi\|_{\mathcal{H}_{\text{loc}}^s},\|\Phi\|_{\mathcal{H}_{\text{loc}}^{s'}}} 1 \quad (20)$$

and

$$\|\phi[t]\|_{\mathcal{H}^{s'}} \lesssim_{s,s',\|\Phi\|_{\mathcal{H}_{\text{loc}}^s},\|\Phi\|_{\mathcal{H}^{s'}}} 1 \quad (21)$$

for all  $s' > 1$ .

*Proof.* By time translation invariance we may take  $t_0 = 0$ . We allow all implied constants to depend on  $s$  and the quantity  $\|\Phi\|_{\mathcal{H}_{\text{loc}}^s}$ , thus for every  $x_0$  there exists  $U_{x_0} \in SO(m, 1)$  such that

$$\|\eta(\cdot - x_0)U_{x_0}(\phi_0)\|_{H_x^s(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U_{x_0}(\phi_1)\|_{H_x^{s-1}(\mathbf{R}^2)} \lesssim 1. \quad (22)$$

In particular, from Sobolev embedding we see that  $U_{x_0}(\phi_0(x_1)) = O(1)$  whenever  $|x_0 - x_1| \leq 1$ . As a consequence, we see that if  $|x_0 - x_1| = O(1)$ , then  $U_{x_0}$  and  $U_{x_1}$  only differ by a left-multiplication by a bounded element of  $SO(m, 1)$ .

For any  $x_0 \in \mathbf{R}^2$ , we can find classical data  $(\phi_{0,x_0}, \phi_{1,x_0})$  with

$$\|\phi_{0,x_0}\|_{H_x^s(\mathbf{R}^2)} + \|\phi_{1,x_0}\|_{H_x^{s-1}(\mathbf{R}^2)} \lesssim 1 \quad (23)$$

which agrees with  $(U_{x_0}(\phi_0), U_{x_0}(\phi_1))$  on  $B(x_0, 1/2)$ ; for instance we can take

$$\tilde{\phi}_0(x_0 + r\omega) := U_{x_0}(\phi_0)(x_0 + f(r)\omega)$$

and

$$\tilde{\phi}_1(x_0 + r\omega) := g(r)U_{x_0}(\phi_1)(x_0 + f(r)\omega)$$

for  $i = 0, 1$ ,  $r \geq 0$ , and  $\omega \in S^1$ , where  $f : [0, +\infty) \rightarrow [0, 1]$  is a smooth function with  $f(r) = r$  for  $r \leq 1/2$  and  $f(r) = 0$  for  $r \geq 1$ , and  $g : [0, +\infty) \rightarrow [0, 1]$  is a smooth function with  $g(r) = 1$  for  $r \leq 1/2$  that vanishes for  $r \geq 1$ ; the bound (23) follows from (22) in the case when  $s \geq 1$  is an integer by direct computation, and the general case follows by an interpolation argument (one can also use the fractional chain rule). For similar reasons we have the more general estimate

$$\begin{aligned} \|\phi_{0,x_0}\|_{H_x^{s'}(\mathbf{R}^2)} + \|\phi_{1,x_0}\|_{H_x^{s'-1}(\mathbf{R}^2)} &\lesssim_{s'} \\ \|\eta(\cdot - x_0)U_{x_0}(\phi_0)\|_{H_x^{s'}(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U_{x_0}(\phi_1)\|_{H_x^{s'-1}(\mathbf{R}^2)} \end{aligned} \quad (24)$$

for all  $s' > 1$ .

We may now apply the local existence theory from<sup>6</sup> [7], and construct a classical local wave map  $(\phi_{x_0}, [-T, T])$  for each fixed  $x_0$  with initial data  $\phi_{x_0}[0] = (U_{x_0}(\phi_{0,x_0}), U_{x_0}(\phi_{1,x_0}))$ , for some  $T \sim 1$  independent of  $x_0$ . Furthermore, by (24) and persistence of regularity theory in [28], we will have

$$\|\phi_{x_0}[t]\|_{H_x^{s'}(\mathbf{R}^2) \times H_x^{s'-1}(\mathbf{R}^2)} \lesssim_{s'} \|\eta(\cdot - x_0)U_{x_0}(\phi_0)\|_{H_x^{s'}(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U_{x_0}(\phi_1)\|_{H_x^{s'-1}(\mathbf{R}^2)} \quad (25)$$

for all  $t \in [-T, T]$ .

By the Lorentz rotation symmetry (1), the rotated solutions  $U_{x_0}^{-1} \circ \phi_{x_0}$  are classical wave maps on  $[-T, T]$  with initial data that agree with  $\Phi$  on  $B(x_0, 1/2)$ . Using finite speed of propagation (and shrinking  $T$  slightly if necessary), we may thus glue all these solutions together to obtain a classical wave map  $\phi$  on  $[-T, T]$  with initial data  $\phi[0] = \Phi$ . The bounds (20), (21) then follow from (25), the stability of Sobolev spaces under multiplication by smooth cutoff functions, and the fact mentioned earlier that  $U_{x_0}$  and  $U_{x_1}$  differ only by bounded rotations when  $|x_1 - x_0| \lesssim 1$ .  $\square$

Standard energy arguments show that classical wave maps are uniquely determined by their initial data. Iterating the above local existence result in the usual fashion, and taking contrapositives, we thus conclude

**Corollary 3.3** (Subcritical blowup criterion). *If  $\Phi = (\phi_0, \phi_1)$  is classical initial data and  $t_0 \in \mathbf{R}$ , then there exists a unique maximal classical lifespan  $I \subset \mathbf{R}$  which is an open interval containing  $t_0$ , and a unique*

---

<sup>6</sup>Strictly speaking, the results in that paper only handle the case when the above norm is sufficiently small, and for a model equation closely related to (15). However, it is not too difficult to adapt the arguments there to the case at hand. Alternatively, one can use the results in [28].

classical wave map  $\phi : I \rightarrow \mathcal{S}$  with  $\phi[t_0] = \Phi$ , (which we call the maximal classical development from the initial data  $\Phi$  at time  $t_0$ ) such that if  $t_*$  is any finite endpoint of  $I$  then  $\|\phi[t]\|_{\mathcal{H}_{\text{loc}}^s} \rightarrow \infty$  as  $t \rightarrow t_*$  for every  $s > 1$ .

*Remark 3.4.* *A posteriori*, we will be able to show that the maximal classical lifespan in Corollary 3.3 is identical to the maximal energy class lifespan in Theorem 1.5 (indeed, this follows immediately from parts (iii) and (v) of that theorem), and so the maximal classical development is essentially the same as the maximal Cauchy development.

#### 4. REDUCTION TO ABSTRACT A PRIORI ESTIMATES

For each compact interval  $I$  and  $E > 0$ , let  $\text{WM}(I, E)$  denote the space of classical wave maps  $\phi : I \rightarrow \mathcal{S}$  on  $I$  of energy less than  $WE$ , quotiented out by the action (1) of the Lorentz rotation group  $SO(m, 1)$ . Note that if  $J \subset I$  then every element of  $\text{WM}(I, E)$  can also be viewed as an element of  $\text{WM}(J, E)$  (indeed, by uniqueness of classical wave maps, we can embed  $\text{WM}(I, E)$  as a subset of  $\text{WM}(J, E)$ ). Also, if  $\phi \in \text{WM}(I, E)$  and  $t \in I$  then  $\phi[t]$  can be viewed as an element of the energy space  $\dot{\mathcal{H}}^1$ . We use  $\text{const} \in \text{WM}(I, E)$  to denote the constant wave map (the exact choice of constant is irrelevant, thanks to Lorentz rotation invariance).

In this section we use standard continuity arguments to reduce Theorem 1.5 to that of establishing a collection of *a priori* estimates for classical wave maps with respect to various “function space norms” (or more precisely, metrics) on  $\text{WM}(I, E)$ . More precisely, we have

**Theorem 4.1** (A priori estimates). *There exist metrics  $d_{S_\mu^1, I}$  on  $\text{WM}(I, E)$  for each compact interval  $I$ ,  $E > 0$  and  $0 < \mu \leq 1$  with the following properties, where we abbreviate  $\|\phi\|_{S_\mu^1(I)}$  for  $d_{S_\mu^1, I}(\phi, \text{const})$ :*

- (i) (*Monotonicity*) If  $I \subset J$ ,  $0 < \mu \leq 1$ , and  $\phi, \phi' \in \text{WM}(J, E)$  then  $d_{S_\mu^1, I}(\phi, \phi') \leq d_{S_\mu^1, J}(\phi, \phi')$ .
- (ii) (*Continuity*) If  $\phi, \phi' \in \text{WM}([t_-, t_+])$  and  $0 < \mu \leq 1$ , then  $d_{S_\mu^1, [a, b]}(\phi, \phi')$  is a continuous function of  $a, b$  in the region  $t_- \leq a < b \leq t_+$ .
- (iii) (*Vanishing*) If  $I_n$  is a decreasing sequence of compact intervals with  $\bigcap_n I_n = \{t_0\}$ ,  $0 < \mu \leq 1$ , and  $\phi, \phi' \in \text{WM}(I_1, E)$ , then  $\lim_{n \rightarrow \infty} d_{S_\mu^1, I_n}(\phi, \phi') \lesssim_E d_{\dot{\mathcal{H}}^1}(\phi[t_0], \phi'[t_0])$ .
- (iv) ( *$S^1$  controls energy*) For any  $I$ , any  $0 < \mu \leq 1$  and  $M > 0$ , any  $\phi, \phi' \in \text{WM}(I, E)$  with  $\|\phi\|_{S_\mu^1(I)}, \|\phi'\|_{S_\mu^1(I)} \leq M$ , and any  $t \in I$ , we have  $d_{\dot{\mathcal{H}}^1}(\phi[t], \phi'[t]) \lesssim_{M, E} d_{S_\mu^1, I}(\phi, \phi')$ .

- (v) (*Persistence of regularity estimate*) If  $M > 0$ ,  $0 < \mu \leq 1$  is sufficiently small depending on  $M$  and  $E$ ,  $I$  is an interval,  $t_0 \in I$ , and  $\phi \in \text{WM}(I, E)$  is such that  $\|\phi\|_{S_\mu^1(I)} \leq M$ , then

$$\|\phi[t]\|_{\mathcal{H}_{\text{loc}}^{1+\delta_0/2}} \lesssim_{M,E,\mu,\|\phi[t_0]\|_{\mathcal{H}_{\text{loc}}^{100}}} 1 \quad (26)$$

for all  $t \in I$ .

- (vi) (*Stability estimate*) If  $M > 0$ ,  $0 < \mu \leq 1$  is sufficiently small depending on  $M$  and  $E$ ,  $I$  is an interval,  $t_0 \in I$ , and  $\phi, \phi' \in \text{WM}(I, E)$  are such that  $\|\phi\|_{S_\mu^1(I)}, \|\phi'\|_{S_\mu^1(I)} \leq M$ , then

$$d_{S_\mu^1, I}(\phi, \phi') \lesssim_{M,\mu,E} d_{\mathcal{H}^1}(\phi[t_0], \phi'[t_0]). \quad (27)$$

*Remark 4.2.* Suppose, as a gross caricature, that wave maps  $\phi$  behaved like scalar solutions  $\phi : I \times \mathbf{R}^3 \rightarrow \mathbf{R}$  to the energy-critical nonlinear wave equation (NLW)  $\square\phi = \phi^5$ . Then one possible choice of the metrics  $d_{S_\mu^1, I}$  would be

$$d_{S_\mu^1, I}(\phi, \phi') := \|\partial_{t,x}(\phi - \phi')\|_{L_t^\infty L_x^2(I \times \mathbf{R}^3)} + \frac{1}{\mu} \|\partial_{t,x}((\phi - \phi')^5)\|_{L_t^5 L_x^{30/13}(I \times \mathbf{R}^3)}.$$

For the small energy theory, it is well known (using Strichartz estimates) that one can iterate in the  $S_\mu^1$  space with  $\mu = 1$  in order to obtain (global) well-posedness for the energy-critical NLW. For large energy, one can similarly iterate<sup>7</sup> in the  $S_\mu^1$  space, with  $\mu$  now chosen sufficiently small depending on the energy, and with the interval  $I$  chosen small enough so that the  $S_\mu^1$  norm stays bounded, to obtain local existence. Thus one should view  $\mu$  as a parameter designed to compensate for large energy. In practice, the Strichartz spaces are insufficient to control the nonlinearities we will encounter, but readers who are familiar with the Strichartz theory for the NLW (or NLS) equations may find the above analogies to be helpful in what follows.

*Remark 4.3.* The persistence of regularity estimate (26) is quite crude. One would expect in fact that the  $\mathcal{H}_{\text{loc}}^s$  norms should grow at most linearly whenever the  $S_\mu^1$  norm is controlled for sufficiently small  $\mu$ , but this seems to require a more delicate analysis and is not actually needed for our applications, so we have elected not to establish this result. Such a claim may however be useful if one eventually wants to obtain scattering for classical wave maps.

We shall prove Theorem 4.1 in later sections. In the remainder of this section, we assume Theorem 4.1 and use it to prove Theorem 1.5. The basic lemma is the following perturbation theory result.

<sup>7</sup>Actually, this is a slight oversimplification. The more precise statement is that boundedness in the  $S_\mu^1$  space allows one to perform a contraction mapping argument in the  $S_1^1$  space.



**Lemma 4.4** (Long-time perturbations). *Let  $I$  be a compact interval let  $E > 0$ , and let  $t_0 \in I$ . Let  $\phi, \phi' \in \text{WM}(I, E)$  be such that*

$$d_{\mathcal{H}^1}(\phi[t_0], \phi'[t_0]) \leq \varepsilon \quad (28)$$

*for some  $\varepsilon > 0$ . Suppose also that one can cover  $I$  by intervals  $I_1, \dots, I_J$  such that*

$$\|\phi\|_{S_\mu^1(I_j)} \leq M \quad (29)$$

*for all  $1 \leq j \leq J$  and some  $M, \mu > 0$ . If  $\mu$  is sufficiently small depending on  $M, E$ , and  $\varepsilon$  is sufficiently small depending on  $M, J, \mu, E$ , then we have*

$$d_{\mathcal{H}^1}(\phi[t], \phi'[t]) \lesssim_{M, E, J, \mu} \varepsilon \quad (30)$$

*and*

$$\|\phi'[t]\|_{\mathcal{H}^{100}} \lesssim_{M, E, J, L, \mu, \|\phi'[t_0]\|_{\mathcal{H}^{100}}} 1. \quad (31)$$

*for all  $t \in I$ , where  $L$  is any upper bound for the length of  $I$ .*

*Proof.* Fix  $E$ ; we allow all constants to depend on  $E$ .

It suffices to verify the claim when  $J = 1$ , since the case of higher  $J$  then follows by iteration (and by energy conservation and Theorem 4.1(i)). By a further subdivision we may then assume that  $t_0$  is either the upper endpoint or lower endpoint of  $I$ . For sake of exposition we shall assume that  $t_0 = \inf(I)$  is the lower endpoint of  $I$ , as the other case is of course exactly analogous. Thus  $I = [t_0, t_1]$  for some  $t_1 > t_0$ .

Our main tool here will be the continuity method. Suppose that  $t_0 < T \leq t_1$  is such that  $\|\phi'\|_{S_\mu^1([t_0, T] \times \mathbf{R}^2)} \leq 3M$ . Then from (27), (28), and Theorem 4.1(i) we see that if  $\mu$  is sufficiently small depending on  $M$ , then

$$d_{S_\mu^1, [t_0, T]}(\phi, \phi') \lesssim_M \varepsilon.$$

Since  $\varepsilon$  is assumed to be sufficiently small depending on  $M$ , we thus conclude from (29) and the triangle inequality that

$$\|\phi'\|_{S_\mu^1([t_0, T] \times \mathbf{R}^2)} \leq 2M.$$

By Theorem 4.1(ii), the set of  $T \in (t_0, t_1]$  such that  $\|\phi'\|_{S_\mu^1([t_0, T] \times \mathbf{R}^2)} \leq 2M$  is thus both closed and open in  $(t_0, t_1]$ . By Theorem 4.1(iii) (if  $M$  is large enough depending on  $E$ ), this set also contains all times sufficiently close to  $t_0$ . As  $(t_0, t_1]$  is connected, we conclude that this set is all of  $(t_0, t_1]$ . In particular we have

$$\|\phi'\|_{S_\mu^1(I)} \leq 2M.$$

Since  $\mu$  is assumed to be sufficiently small depending on  $M$ , we can apply Theorem 4.1(v) and (28) to conclude that

$$d_{S^1, I}(\phi, \phi') \lesssim_{M, \mu} \varepsilon.$$

The claim (30) then follows from Theorem 4.1(iv). To prove (31), observe from Theorem 4.1(v) that

$$\sup_{t \in I} \|\phi'[t]\|_{\mathcal{H}_{\text{loc}}^{1+\delta_0/2}} \lesssim_{M, \mu, \|\phi'[t_0]\|_{\mathcal{H}^{100}}} 1,$$

and the claim then follows from repeated application of Lemma 3.2.  $\square$

**Corollary 4.5** (Lifespan stability). *Let  $I$  be a compact interval, let  $t_0 \in I$  and  $E > 0$ , and let  $\phi \in \text{WM}(I, E)$ . Suppose that one can cover  $I$  by intervals  $I_1, \dots, I_J$  such that  $\|\phi\|_{S_\mu^1(I_j)} \leq M$  for all  $1 \leq j \leq J$  and some  $M, \mu > 0$ . Let  $\Phi' \in \mathcal{S}$  be such that  $d_{\mathcal{H}^1}(\phi[t_0], \Phi') \leq \varepsilon$  for some  $\varepsilon > 0$ . If  $\mu$  is sufficiently small depending on  $M$  and  $E$ , and  $\varepsilon$  is sufficiently small depending on  $M, J, \mu$ , then there exists  $\phi' \in \text{WM}(I, E)$  such that  $\phi'[t_0] = \Phi'$ .*

*Proof.* If the conclusion failed, then by Corollary 3.3 there would exist an open subinterval  $I' \subset I$  containing  $t_0$  and a classical solution  $\phi' : I' \rightarrow \mathcal{S}$  with  $\phi'[t_0] = \Phi'$  such that  $\|\phi'[t]\|_{\mathcal{H}_{\text{loc}}^{1+\delta_0/2}}$  (and thus  $\|\phi'[t]\|_{\mathcal{H}^{100}}$ ) was unbounded on  $I'$ . But this contradicts Lemma 4.4 (restricted to  $I'$ ).  $\square$

We are now ready to prove Theorem 1.5, which will be a completely soft argument relying on Lemma 4.4, Corollary 4.5, and gluing of intervals.

*Proof of Theorem 1.5.* We first observe that it suffices to establish the claim under the additional assumption that all solutions involved have energy less than an arbitrary parameter  $E > 0$ , since the general case then follows by letting  $E \rightarrow \infty$  and using the various claims in Theorem 1.5 to ensure that the solutions constructed for different choices of  $E$  are compatible with each other.

We now fix this energy threshold  $E > 0$ , choose a sufficiently large parameter  $M > 0$  depending on  $E$ , and also choose a sufficiently small parameter  $\mu$  depending on  $E, M$ . We allow all implied constants to depend on  $E$ .

Given  $\Phi_0 \in \dot{\mathcal{H}}^1$  with  $E(\Phi_0) < E$  and  $t_0 \in \mathbf{R}$ , let us say that a compact interval  $I$  containing  $t_0$  is *good* for this choice of initial data if there exists a sequence  $\phi^{(n)} \in \text{WM}(I, E)$  such that  $\phi^{(n)}(t_0)$  converges to  $\Phi_0$  in  $\dot{\mathcal{H}}^1$  (and in particular has energy less than  $E$  for sufficiently large  $n$ , by Theorem 1.2(iii)), and such that  $I$  can be covered by finitely many compact intervals  $I_1, \dots, I_J$ , such that  $\limsup_{n \rightarrow \infty} \|\phi^{(n)}\|_{S_\mu^1(I_j)} < M$  for each  $1 \leq j \leq J$ . From Lemma 4.4, we see (for appropriate choices of  $M, \mu$ ) that  $\phi^{(n)}$  forms a Cauchy sequence in the uniform topology in  $\dot{\mathcal{H}}^1$  on  $I$ , and so (by completeness of  $\dot{\mathcal{H}}^1$ ) converges uniformly in  $\dot{\mathcal{H}}^1$  to a

limit  $\phi : I \rightarrow \dot{\mathcal{H}}^1$ , which is then continuous in  $\dot{\mathcal{H}}^1$  since each of the  $\phi^{(n)}$ , being classical, are automatically continuous in  $\dot{\mathcal{H}}^1$  by Theorem 1.2. Another application of Lemma 4.4 shows that this limit is independent of the choice of sequence  $\phi^{(n)}$ . Similarly, if there are two good compact intervals  $I, I'$  containing  $t_0$ , and  $\phi : I \rightarrow \dot{\mathcal{H}}^1$  and  $\phi' : I' \rightarrow \dot{\mathcal{H}}^1$  are constructed as above from the same initial data, then another application of Lemma 4.4 (applied to the common intersection  $I \cap I'$  and using monotonicity) shows that  $\phi$  and  $\phi'$  agree on  $I \cap I'$ . Finally, another application of Lemma 4.4 shows that the union of finitely many good intervals is still good. Thus, if we define the *maximal lifespan*  $I$  of  $\Phi_0$  from  $t_0$  to be the union of all the good compact intervals, we can define a unique maximal Cauchy development  $\phi : I \rightarrow \dot{\mathcal{H}}^1$  by gluing together all the partial developments  $\phi : J \rightarrow \dot{\mathcal{H}}^1$  associated to good intervals.

We now establish claim (i), namely that  $I$  is an open interval containing  $t_0$ . We first show that  $I$  contains a neighbourhood of  $t_0$ . By Theorem 1.2(i), there exists a sequence  $\Phi^{(n)} \in \mathcal{S}$  that converges in  $\dot{\mathcal{H}}^1$  to  $\Phi$ , and thus (by Theorem 1.2(iii)) we have  $E(\Phi^{(n)}) < E$  after discarding at most finitely many  $n$ . Let  $\varepsilon > 0$  be sufficiently small depending on  $E, M, \mu$ . We find an  $n$  such that  $d_{\dot{\mathcal{H}}^1}(\Phi^{(n)}, \Phi) \leq \varepsilon/2$ , and then let  $\phi^{(n)} : I^{(n)} \rightarrow \mathcal{S}$  be the maximal classical development of  $\Phi^{(n)}$  from  $t_0$  given by Lemma 3.3, thus  $I^{(n)}$  is an open interval containing  $t_0$ . Applying Theorem 4.1(iii), we can find a compact interval  $I' \subset I^{(n)}$  containing  $t_0$  in its interior such that  $\limsup_{n \rightarrow \infty} \|\phi^{(n)}\|_{S_\mu^1(I')} < M/2$ , if  $M$  is large enough depending on  $E$ . Now for all sufficiently large  $m$ , we have  $d_{\dot{\mathcal{H}}^1}(\Phi^{(n)}, \Phi^{(m)}) \leq \varepsilon$  by the triangle inequality, so by Corollary (4.5) we have a classical wave map  $\phi^{(m)} : I' \rightarrow \mathcal{S}$  with  $\phi^{(m)}[t_0] = \Phi^{(m)}$ . Inspecting the proof of Lemma 4.4 we see that  $\limsup_{m \rightarrow \infty} \|\phi^{(m)}\|_{S_\mu^1(I')} < M$ , and so  $I'$  is good. Thus  $I$  contains an open neighbourhood of  $t_0$ .

It remains to show that  $I$  itself is open. Let  $t_1 \in I$ , then by construction there exists a good interval  $I'$  containing both  $t_0$  and  $t_1$ . Thus we can find a sequence  $\phi^{(n)} \in \text{WM}(J, E)$  which converges uniformly to  $\phi$  in  $\dot{\mathcal{H}}^1$  on  $I'$ , in particular  $\phi^{(n)}[t_1]$  converges to  $\phi[t_1]$  in  $\dot{\mathcal{H}}^1$ . Applying the previous argument at time  $t_1$  rather than  $t_0$ , we conclude that there exists a compact interval  $K$  containing  $t_1$  in its interior such that the maximal classical developments of  $\phi^{(n)}$  exist on  $K$  and  $\limsup_{n \rightarrow \infty} \|\phi^{(n)}\|_{S_\mu^1(K)} < M$ . This implies that  $J \cup K$  is good, and so  $t_1$  is in the interior of  $I$ . Thus  $I$  is open as claimed.

The claim (ii) of Theorem 1.5 is clear from construction. Claim (iii) follows easily from Corollary 4.5 (and the uniqueness of classical wave maps), so we now turn to claim (iv). From the proof of (i) we know that every element of the maximal lifespan of a solution is contained

in the interior of a good interval, and so by compactness, we see that to prove (iv), it suffices to do so in the case when  $K$  is a good interval. Let  $\phi_\infty : K \rightarrow \mathcal{H}^1$  be the restriction of a maximal Cauchy development to a good interval  $K$ , with energy strictly less than  $E$ . Let  $\phi_\infty^{(m)} \in \text{WM}(K)$  be a sequence converging uniformly to  $\phi_\infty$  as in the definition of a good interval, thus  $E(\phi_\infty^{(m)}) < E$  for sufficiently large  $m$ , and one can cover  $K$  by a finite number of intervals  $K_1, \dots, K_k$  such that  $\limsup_{m \rightarrow \infty} \|\phi_\infty^{(m)}\|_{S_\mu^1(K_j)} < M$  for all  $1 \leq j \leq k$ . Now let  $\varepsilon > 0$  be a sufficiently small quantity depending on  $E, M, k, \mu$ . For  $n$  large enough,  $\Phi_{0,n}$  lies within  $\varepsilon/3$  of  $\Phi_{0,\infty}$  in the energy metric and has energy less than  $E$ , and so if one writes  $\Phi_{0,n}$  as the limit of smooth data  $\Phi_{0,n}^{(m)}$  of energy less than  $E$ , we see that  $\Phi_{0,n}^{(m)}$  lies within  $\varepsilon$  of  $\phi_\infty^{(m')}[t_0]$  if  $m, m'$  are large enough. Applying Corollary 4.5 and Lemma 4.4 one can then (if  $\varepsilon$  is small enough) find  $\phi_n^{(m)} \in \text{WM}(K)$  with  $\phi_n^{(m)}[t_0] = \Phi_{0,n}^{(m)}$ , and furthermore (by inspection of the proof of Lemma 4.4) we have  $\limsup_{m \rightarrow \infty} \|\phi_n^{(m)}\|_{S_\mu^1(I)} < M$  on  $O_{M,k,K}(1)$  intervals  $I$  (depending on  $m'$ , but independent of  $m$ ) covering  $K$ . From this we see that  $K$  is good for  $\Phi_n$ , and from Lemma 4.4 we then see that the resulting Cauchy development  $\phi_n : K \rightarrow \mathcal{H}^1$  differs from  $\phi_\infty : K \rightarrow \mathcal{H}^1$  by  $O_{M,k,K}(\varepsilon)$  in the uniform energy metric. Sending  $\varepsilon \rightarrow 0$  we conclude that  $\phi_n$  converges uniformly to  $\phi_\infty$  on  $K$ , and Claim (iv) follows.

Finally, we establish Claim (v). Suppose for contradiction that there was a sequence  $t_n \in I$  converging to a finite endpoint  $t_* \in \mathbf{R}$  of  $I$  such that  $\phi[t_n]$  converged in  $\mathcal{H}^1$  to a limit  $\Phi_*$ . Since  $\phi$  had energy less than  $E$ , we see that  $\Phi_*$  does too. By the arguments used to prove (i), we can find a compact interval  $J$  containing  $t_*$  in its interior for which there exists a sequence  $\phi_*^{(m)} \in \text{WM}(J, E)$  with  $\limsup_{m \rightarrow \infty} \|\phi_*^{(m)}\|_{S_\mu^1(J)} < M$  and with  $\phi_*^{(m)}(t_*)$  converging in  $\mathcal{H}^1$  to  $\Phi_*$  (and thus has energy less than  $E$  for sufficiently large  $m$ ). As before,  $\phi_*^{(m)}$  converges uniformly in  $\mathcal{H}^1$  to a continuous limit  $\phi_* : J \rightarrow \mathcal{H}^1$ . Since  $t_n \in J$  for sufficiently large  $n$ , we see from continuity that  $\phi_*[t_n]$  converges in  $\mathcal{H}^1$  to  $\Phi_*$ , and thus by the triangle inequality  $d_{\mathcal{H}^1}(\phi_*[t_n], \phi[t_n])$  converges to zero.

Let  $\varepsilon$  be sufficiently small depending on  $M, E$  and  $\mu$ . Then for  $n$  large enough, we have  $d_{\mathcal{H}^1}(\phi_*[t_n], \phi[t_n]) \leq \varepsilon/2$ , and thus if  $\Phi_n^{(m)} \in \mathcal{S}$  is a sequence of classical data converging in  $\mathcal{H}^1$  to  $\phi[t_n]$  (and thus with energy strictly less than  $E$  for sufficiently large  $n$ ) then  $d_{\mathcal{H}^1}(\phi_*^{(m')}[t_n], \Phi_n^{(m)}) \leq \varepsilon$  whenever  $m, m'$  are sufficiently large. Applying Corollary 4.5 and Lemma 4.4, we see (if  $\varepsilon$  is small enough) that there is  $\phi_n^{(m)} \in \text{WM}(J, E)$  with  $\phi_n^{(m)}[t_n] = \Phi_n^{(m)}$  such that  $\limsup_{m \rightarrow \infty} \|\phi_n^{(m)}\|_{S_\mu^1(J)} < M$ . From this (and the uniqueness of classical wave maps) we see that if  $K$  is any good

interval for  $\phi$  that contains both  $t_0$  and  $t_n$ , then  $K \cup J$  is also good for  $\phi$ . Thus  $J \subset I$ , and therefore  $t_*$  is in fact an interior point of  $I$ , a contradiction. This establishes Claim (v), and the proof of Theorem 1.5 is complete.  $\square$

It remains to prove Theorem 4.1. This is the purpose of the remaining sections of this paper.

## 5. THE CALORIC GAUGE AND THE ENERGY SPACE

Our arguments thus far have been extremely abstract; the exact form of the wave maps equation, or even of the energy space  $\dot{\mathcal{H}}^1$ , has not played any real role in previous sections, beyond such basic properties as energy conservation and classical well-posedness. To proceed further, we will have to use the specific structure of the equation and the energy space. In this section we recall from [25] the construction of the energy space, and in particular the *caloric gauge* that is used in that construction. We refer the reader to [25] for a more detailed treatment of the material here.

**5.1. The caloric gauge.** We first recall a global existence theorem for the harmonic map heat flow, essentially due to Eells and Sampson [3].

**Theorem 5.2** (Global existence for heat flow). *Let  $I$  be a compact interval, and let  $\phi : I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  be a smooth map which differs from a constant  $\phi(\infty)$  by a Schwartz function. Then there exists a unique extension  $\phi : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  of  $\phi$  (thus  $\phi(0, t, x) = \phi(t, x)$  for all  $(t, x) \in I \times \mathbf{R}^2$ ) which is smooth with all derivatives bounded, and obeys the harmonic map heat flow equation*

$$\partial_s \phi = (\phi^* \nabla)_i \partial_i \phi$$

*on  $\mathbf{R}^+ \times I \times \mathbf{R}^2 := \{(s, t, x) : s \in \mathbf{R}^+, t \in I, x \in \mathbf{R}^2\}$ , and converges in  $C_{\text{loc}}^\infty(I \times \mathbf{R}^2)$  to  $\phi(\infty)$  as  $s \rightarrow \infty$ .*

*Proof.* See [25, Theorem 3.16].  $\square$

This result allows us to extend any classical wave map  $\phi \in WM(I)$  by harmonic map heat flow into the region  $\mathbf{R}^+ \times I \times \mathbf{R}^2$ . In order to analyse this heat flow we will place an orthonormal frame on this map (or more precisely on the pullback tangent bundle  $\phi^*T\mathbf{H}$ ).

Given any point  $p \in \mathbf{H}$ , define an *orthonormal frame* at  $p$  to be any orthogonal orientation-preserving map  $e : \mathbf{R}^m \rightarrow T_p\mathbf{H}$  from  $\mathbf{R}^m$

to the tangent space at  $p$  (with the metric  $h(p)$ , of course), and let  $\text{Frame}(T_p\mathbf{H})$  denote the space of such frames; note that this space has an obvious transitive action of the special orthogonal group  $SO(m)$ . We then define the *orthonormal frame bundle*  $\text{Frame}(\phi^*T\mathbf{H})$  of  $\phi$  to be the space of all pairs  $((s, t, x), e)$  where  $(s, x) \in \mathbf{R}^+ \times I \times \mathbf{R}^2$  and  $e \in \text{Frame}(T_{\phi(s, x)}\mathbf{H})$ ; this is a smooth vector bundle over  $\mathbf{R}^+ \times I \times \mathbf{R}^2$ . We then define an *orthonormal frame*  $e \in \Gamma(\text{Frame}(\phi^*T\mathbf{H}))$  for  $\phi$  to be a section of this bundle, i.e. a smooth assignment  $e(s, x) \in \text{Frame}(T_{\phi(s, x)}\mathbf{H})$  of an orthonormal frame at  $\phi(s, x)$  to every point  $(s, x) \in \mathbf{R}^+ \times I \times \mathbf{R}^2$ .

Each orthonormal frame  $e \in \Gamma(\text{Frame}(\phi^*T\mathbf{H}))$  provides an orthogonal, orientation-preserving identification between the vector bundle  $\phi^*T\mathbf{H}$  (with the metric  $\phi^*h$ ) and the trivial bundle  $(\mathbf{R}^+ \times I \times \mathbf{R}^2) \times \mathbf{R}^m$  (with the Euclidean metric on  $\mathbf{R}^m$ ), thus sections  $\Psi \in \Gamma(\phi^*T\mathbf{H})$  can be pulled back to functions  $e^*\Psi : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \mathbf{R}^m$  by the formula  $e^*\Psi := e^{-1} \circ \Psi$ . The connection  $\phi^*\nabla$  on  $\phi^*T\mathbf{H}$  can similarly be pulled back to a connection  $D$  on the trivial bundle  $(\mathbf{R}^+ \times \mathbf{R}^2) \times \mathbf{R}^m$ , defined by

$$D_i := \partial_i + A_i \quad (32)$$

where  $A_i \in \mathfrak{so}(m)$  is the skew-adjoint  $m \times m$  matrix field is given by the formula

$$(A_i)_{ab} = \langle (\phi^*\nabla)_i e_a, e_b \rangle_{\phi^*h} \quad (33)$$

where  $e_1, \dots, e_m$  are the images of the standard orthonormal basis for  $\mathbf{R}^m$  under  $e$ . Of course one similarly has covariant derivatives  $D_t = \partial_t + A_t$ ,  $D_s = \partial_s + A_s$  in the  $t$  and  $s$  directions.

Given such a frame, we define the *derivative fields*  $\psi_j : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \mathbf{R}^m$  by the formula

$$\psi_j := e^*\partial_j\phi, \quad (34)$$

and similarly define

$$\psi_s := e^*\partial_s\phi; \quad \psi_t := e^*\partial_t\phi. \quad (35)$$

We record the zero-torsion property

$$D_i\psi_j = D_j\psi_i \quad (36)$$

and the constant negative curvature property

$$[D_i, D_j] = \partial_i A_j - \partial_j A_i + [A_i, A_j] = -\psi_i \wedge \psi_j \quad (37)$$

where  $\psi_i \wedge \psi_j$  is the anti-symmetric matrix field

$$\psi_i \wedge \psi_j := \psi_i \psi_j^* - \psi_j \psi_i^*.$$

The harmonic map heat flow equation becomes

$$\psi_s = D_i\psi_i = \partial_i\psi_i + A_i\psi_i. \quad (38)$$

The wave maps equation (3) becomes

$$w|_{s=0} = 0 \quad (39)$$

where  $w : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \mathbf{R}^m$  is the *wave-tension field*

$$w := D^\alpha \psi_\alpha = \partial^\alpha \psi_\alpha + A^\alpha \psi_\alpha. \quad (40)$$

Following [25], we say that an orthonormal frame  $e$  is a *caloric gauge* if one has

$$A_s = 0 \quad (41)$$

throughout  $\mathbf{R}^+ \times I \times \mathbf{R}^2$ , and if we have

$$\lim_{s \rightarrow \infty} e(s, t, x) = e(\infty) \quad (42)$$

for all  $x \in \mathbf{R}^2$  and some constant  $e(\infty) \in \text{Frame}(T_{\phi(\infty)}\mathbf{H})$ .

We have the following existence theorem for this gauge:

**Theorem 5.3** (Existence of caloric gauge). *Let  $I$  be a time interval with non-empty interior, let  $\phi : I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  be a smooth map differing from a constant  $\phi(\infty)$  by a Schwartz function, and let  $e(\infty) \in \text{Frame}(T_{\phi(\infty)}\mathbf{H})$  be a frame for  $\phi(\infty)$ . Let  $\phi : \mathbf{R}^+ \times I \times \mathbf{R}^2 \rightarrow \mathbf{H}$  be the heat flow extension from Theorem 5.2. Then there exists a unique smooth frame  $e \in \Gamma(\text{Frame}(\phi^*T\mathbf{H}))$  such that  $e(t)$  is a caloric gauge for  $\phi(t)$  which equals  $e(\infty)$  at infinity for each  $t \in I$ . All derivatives of  $\phi - \phi(\infty)$  in the variables  $t, x, s$  are Schwartz in  $x$  for each fixed  $t, s$ . In particular, these derivatives are uniformly Schwartz for  $t, s$  in a compact range.*

*Proof.* See [25, Theorem 3.16]. □

In the caloric gauge we have the derivative fields  $\psi_t, \psi_x, \psi_s$  and the connection fields  $A_t, A_x$  (recall in the caloric gauge that the field  $A_s$  is trivial). It will be convenient to introduce the combined vector or tensor fields

$$\begin{aligned} \Psi_x &:= (\psi_x, A_x) \\ \psi_{t,x} &:= (\psi_t, \psi_x) \\ A_{t,x} &:= (A_t, A_x) \\ \Psi_{t,x} &:= (\psi_{t,x}, A_{t,x}) \\ \Psi_{s,t,x} &:= (\psi_s, \psi_{t,x}, A_{t,x}). \end{aligned}$$

We refer to  $\Psi_{s,t,x}$  as the *differentiated fields* of  $\phi$  in the caloric gauge associated with  $e(\infty)$ . We also introduce the *wave-tension field*  $w := D^\alpha \psi_\alpha$ ; the wave map equation (3) is then equivalent to the vanishing of this quantity on the  $s = 0$  boundary of  $\mathbf{R}^+ \times I \times \mathbf{R}^2$ . We recall the basic equations of motion for these fields:

**Proposition 5.4** (Equations of motion). *Let  $I$  be an interval, let  $\phi \in \text{WM}(I, E)$  with energy  $E(\phi) \leq E$ , let  $e$  be a caloric gauge for  $\phi$ , and let  $\phi_t, \psi_x, \psi_s, A_t, A_x, w$  be the differentiated fields, connection fields, and wave-tension field. Then we have the equations of motion*

$$A_{t,x}(s, t, x) = \int_s^\infty \psi_s \wedge \psi_{t,x}(s', t, x) \, ds' \quad (43)$$

$$\psi_{t,x}(s, t, x) = - \int_s^\infty D_{t,x} \psi_s(s', t, x) \, ds' \quad (44)$$

$$\partial_s \psi_s = D_i D_i \psi_s - (\psi_s \wedge \psi_i) \psi_i \quad (45)$$

$$\partial_s \psi_{t,x} = D_i D_i \psi_{t,x} - (\psi_{t,x} \wedge \psi_i) \psi_i \quad (46)$$

$$\partial_s w = D_i D_i w - (w \wedge \psi_i) \psi_i + 2(\psi_\alpha \wedge \psi_i) D_i \psi^\alpha \quad (47)$$

$$D^\alpha D_\alpha \psi_s = \partial_s w - (\psi_\alpha \wedge \psi_s) \psi^\alpha. \quad (48)$$

*Proof.* The equations (43), (44), (45), (46) follow from [25, Lemma 3.11] and the qualitative decay properties at  $s = \infty$  from [25, Theorem 3.16]. The equation (47) follows from [25, Lemma 7.1] (noting from (36) that  $(\psi_\alpha \wedge \psi_i) D_i \psi^\alpha = -(\psi_t \wedge \psi_i) D_t \psi_i$ ). Finally, (48) follows from (36), (37), and the definition of  $w$ .  $\square$

*Remark 5.5.* As a caricature, in which one pretends  $\phi$  is a scalar field rather than taking values in  $\mathbf{H}$ , the reader is encouraged to use the heuristics  $\psi_\alpha(s, t) \approx \partial_\alpha e^{s\Delta} \phi(t)$ , and  $\psi_s(s, t) \approx \Delta e^{s\Delta} \phi(t)$ . Combining these heuristics with (43), we see that  $A_\alpha$  should behave like a para-product of  $\phi$  with itself; see [21] or [22, Chapter 6] for further discussion.

*Remark 5.6.* The six evolution equations (43)-(48) will be used in different ways. The equations (43), (44) effectively express the fields  $A_\alpha, \psi_\alpha$  in terms of  $\psi_s$ , thus in principle converting the wave maps equation into a “scalar” equation involving  $\psi_s$  as the only “dynamic” field (cf. the “dynamic separation” used in [10]). We will evolve this field by the wave equation (48), using the parabolic equation (47) to solve for the  $\partial_s w$  forcing term appearing in that wave equation. Finally, the remaining equations (45), (46) will be used to compute the initial data  $\psi_s(s, 0, x)$ ,  $D_t \psi_s(s, 0, t) = \partial_s \psi_t(s, 0, t)$  of the wave equation (48).

*Remark 5.7.* It will be convenient to unify (43), (44) in the schematic form

$$\Psi_{t,x}(s) = \int_s^\infty \mathcal{O}(\partial_{t,x} \psi_s(s')) + \mathcal{O}(\psi_s(s') \Psi_{t,x}(s')) \, ds'. \quad (49)$$

Similarly we have

$$\Psi_x(s) = \int_s^\infty \mathcal{O}(\partial_x \psi_s(s')) + \mathcal{O}(\psi_s(s') \Psi_x(s')) \, ds'. \quad (50)$$

These integral ODE thus, in principle, recover  $\Psi_x$  or  $\Psi_{t,x}$  from  $\psi_s$  (assuming sufficient decay at  $s = \infty$ ). In a similar spirit, we have the



identities

$$\partial_s \Psi_{t,x} = \mathcal{O}(\partial_{t,x} \psi_s) + \mathcal{O}(\psi_s \Psi_{t,x}) \quad (51)$$

$$\partial_{t,x} \psi_s = \mathcal{O}(\partial_s \Psi_{t,x}) + \mathcal{O}(\psi_s \Psi_{t,x}) \quad (52)$$

$$\partial_{t,x} \Psi_x = \partial_x \Psi_{t,x} + \mathcal{O}(\Psi_x \Psi_{t,x}) \quad (53)$$

$$\psi_s = \mathcal{O}(\partial_x \Psi_x) + \mathcal{O}(\Psi_x^2) \quad (54)$$

arising from (36), (37), (38), and similarly with the  $t$  subscripts dropped.

**5.8. Working in the caloric gauge.** Let  $\phi \in \text{WM}(I, E)$  be a classical wave map on a compact time interval  $I$ , quotiented out by Lorentz rotations  $SO(m, 1)$ . By Theorem 5.3, given any such wave map, and given any frame  $e(\infty)$  at  $e(\infty) \in \text{Frame}(T_{\phi(\infty)} \mathbf{H})$ , one can construct a caloric gauge for  $\phi$ , which then generates the differentiated fields  $\Psi_{s,t,x}$ . We let  $\text{WMC}(\phi, I)$  denote the collection of all the possible fields  $\Psi_{s,t,x}$  that can arise for a fixed  $\phi$  by choosing  $e(\infty)$ , and let  $\text{WMC}(I, E) = \bigcup_{\phi \in \text{WM}(I, E)} \text{WMC}(\phi, I)$  denote all the differentiated fields that can arise from some classical wave map  $\phi \in \text{WM}(I, E)$ .

Observe that if  $U \in SO(m)$  is a rotation, then replacing  $e(\infty)$  by  $e(\infty) \circ U^{-1}$  corresponds to rotating the fields  $\psi_{t,x}$  to  $U\psi_{t,x}$ , and similarly rotating  $A_{t,x}$  to  $UA_{t,x}U^{-1}$ . This gives an action of  $SO(m)$  on  $\text{WMC}(I, E)$ , whose orbits are precisely the sets  $\text{WMC}(\phi, I)$ . Thus each classical wave map  $\phi \in \text{WM}(I, E)$  gives rise to an  $SO(m)$ -orbit  $\text{WMC}(\phi, I)$  of differentiated fields. Conversely, it is not hard to see (from the Picard uniqueness theorem) that each differentiated field  $\Psi_{s,t,x} \in \text{WMC}(I, E)$  arises from exactly one wave map  $\phi \in \text{WM}(I, E)$  (quotienting out by  $SO(m, 1)$  as usual). So we have an identification

$$\text{WM}(I, E) \equiv SO(m) \backslash \text{WMC}(I, E).$$

We can play similar games with the initial data space  $\mathcal{S}$ : if  $\Phi = (\phi_0, \phi_1) \in \mathcal{S}$  is classical initial data, we can construct associated caloric gauges  $\Psi_{s,t,x}$  on  $\mathbf{R}^+ \times \mathbf{R}^2$  for each choice of frame  $\mathbf{e}(\infty) \in \text{Frame}(T_{\phi_0(\infty)} \mathbf{H})$ , for instance by extending  $\Phi$  for a short amount of time and then using Theorem 5.3, or by using [25, Theorem 3.12], [25, Lemma 4.8]. We let  $\mathcal{SC}(\Phi)$  denote the collection of all such differentiated fields for a fixed  $\Phi$ , and  $\mathcal{SC} := \bigcup_{\Phi \in \mathcal{S}} \mathcal{SC}(\Phi)$  denote the total collection of such fields. Once again, the orthogonal group  $SO(m)$  acts on  $\mathcal{SC}$ , with orbits  $\mathcal{SC}(\Phi)$ , with the orbit determining  $\Phi$  up to the action of  $SO(m, 1)$ , so that we have an identification

$$SO(m, 1) \backslash \mathcal{S} \equiv SO(m) \backslash \mathcal{SC}.$$

Also observe that if  $\Psi_{s,t,x} \in \text{WMC}(\phi, I)$  for some  $\phi \in \text{WM}(I, E)$ , then  $\Psi_{s,t,x}(t_0) \in \mathcal{SC}(\phi[t_0])$  for all  $t_0 \in I$ . Finally, the constant data

$\text{const} \in \mathcal{S}$  corresponds to the zero differentiated field  $0 \in \mathcal{SC}$ , thus  $\mathcal{SC}(\text{const}) = \{0\}$  (and similarly  $\text{WMC}(\text{const}) = \{0\}$ ).

**5.9. The energy space.** We introduce the Littlewood-Paley space

$$\mathcal{L} := L^2(\mathbf{R}^+ \times \mathbf{R}^2 \rightarrow \mathbf{R}^m, dx ds) \times L^2(\mathbf{R}^2 \rightarrow \mathbf{R}^m, \frac{1}{2} dx). \quad (55)$$

We can define an energy metric  $d_{\mathcal{H}^1}$  on  $\mathcal{SC}$  by the formula

$$\begin{aligned} d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x}) &:= \|(\psi_s - \psi'_s, \psi_t(0) - \psi'_t(0))\|_{\mathcal{L}} \\ &= \left( \int_0^\infty \|\delta\psi_s(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds + \frac{1}{2} \|\delta\psi_t(0)\|_{L_x^2(\mathbf{R}^2)}^2 \right)^{1/2}. \end{aligned} \quad (56)$$

We also abbreviate  $d_{\mathcal{H}^1}(\Psi_{s,t,x}, 0)^2$  as  $E(\Psi_{s,t,x})$ , and refer to this as the *energy* of the differentiated field  $\Psi_{s,t,x}$ .

Observe that this metric is preserved by the action of  $SO(m)$ . It thus descends to a metric on  $SO(m, 1) \backslash \mathcal{S} \equiv SO(m) \backslash \mathcal{SC}$  in the usual manner; by abuse of notation we also denote this metric as  $d_{\mathcal{H}^1}$ , thus

$$d_{\mathcal{H}^1}(\Phi, \Phi') := \inf_{\Psi_{s,t,x} \in \mathcal{SC}(\Phi)} \inf_{\Psi'_{s,t,x} \in \mathcal{SC}(\Phi')} d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x}).$$

We define the energy space<sup>8</sup>  $\dot{\mathcal{H}}^1$  to be the completion of  $SO(m, 1) \backslash \mathcal{S}$  using the metric  $\dot{\mathcal{H}}^1$ , and  $\iota$  to be the quotient map from  $\mathcal{S}$  to  $SO(m, 1) \backslash \mathcal{S}$ . With these definitions, Theorem 1.2 was established in [25]. Observe from Theorem 1.2(iv) that the energy of a map  $\Phi \in \mathcal{S}$  is equal to the energy of any of its differentiated fields  $\Psi_{s,t,x} \in \mathcal{SC}(\Phi)$ .

**5.10. Lifting Theorem 4.1 to the caloric gauge.** We will be working exclusively in the caloric gauge in order to prove Theorem 4.1. Because of this, it will be convenient to deduce Theorem 4.1 from a variant involving differentiated wave maps  $\text{WMC}(I, E)$  in the caloric gauge, rather than classical wave maps  $\text{WM}(I, E)$ . More precisely we deduce Theorem 4.1 from

**Theorem 5.11** (A priori estimates in the caloric gauge). *There exist metrics  $d_{S_\mu^1, I}$  on  $\text{WMC}(I, E)$  for each compact interval  $I$ ,  $E > 0$ , and  $0 < \mu \leq 1$  with the following properties, where we abbreviate  $\|\Psi_{s,t,x}\|_{S_\mu^1(I)}$  for  $d_{S_\mu^1, I}(\Psi_{s,t,x}, 0)$ :*

- (i) (*Monotonicity*) If  $I \subset J$ ,  $0 < \mu \leq 1$ , and  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WM}(J, E)$  then  $d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) \leq d_{S_\mu^1, J}(\Psi_{s,t,x}, \Psi'_{s,t,x})$ .

---

<sup>8</sup>This definition is equivalent to that in [25], though arranged slightly differently.

- (ii) (*Continuity*) If  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}([t_-, t_+])$  and  $0 < \mu \leq 1$ , then  $d_{S_\mu^1, [a,b]}(\Psi_{s,t,x}, \Psi'_{s,t,x})$  is a continuous function of  $a, b$  in the region  $t_- \leq a < b \leq t_+$ .
- (iii) (*Vanishing*) If  $I_n$  is a decreasing sequence of compact intervals with  $\bigcap_n I_n = \{t_0\}$ ,  $0 < \mu \leq 1$ , and  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I_1)$ , then  $\lim_{n \rightarrow \infty} d_{S_\mu^1, I_n}(\Psi_{s,t,x}, \Psi'_{s,t,x}) \lesssim_E d_{\mathcal{H}^1}(\Psi_{s,t,x}(t_0), \Psi'_{s,t,x}(t_0))$ .
- (iv) ( *$S^1$  controls energy*) For any  $I$ , any  $0 < \mu \leq 1$  and  $M > 0$ , any  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$  with  $\|\Psi_{s,t,x}\|_{S_\mu^1(I)}, \|\Psi'_{s,t,x}\|_{S_\mu^1(I)} \leq M$ , and any  $t \in I$ , we have  $d_{\mathcal{H}^1}(\Psi_{s,t,x}(t), \Psi'_{s,t,x}(t)) \lesssim_{M,E} d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x})$ .
- (v) (*Persistence of regularity estimate*) If  $M > 0$ ,  $0 < \mu \leq 1$  is sufficiently small depending on  $M$  and  $E$ ,  $I$  is an interval,  $t_0 \in I$ , and  $\phi \in \text{WM}(I, E)$  is such that  $\|\Psi_{s,t,x}\|_{S_\mu^1(I)} \leq M$  for some  $\Psi_{s,t,x} \in \text{WMC}(\phi, I)$ , then

$$\|\phi[t]\|_{\mathcal{H}_{\text{loc}}^{1+\delta_0/2}} \lesssim_{M,E,\mu,\|\phi[t_0]\|_{\mathcal{H}_{\text{loc}}^{100}}} 1 \quad (57)$$

for all  $t \in I$ .

- (vi) (*Stability estimate*) If  $M > 0$ ,  $0 < \mu \leq 1$  is sufficiently small depending on  $M$  and  $E$ ,  $I$  is an interval,  $t_0 \in I$ , and  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WM}(I, E)$  are such that  $\|\Psi_{s,t,x}\|_{S_\mu^1(I)}, \|\Psi'_{s,t,x}\|_{S_\mu^1(I)} \leq M$ , then

$$d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) \lesssim_{M,\mu,E} d_{\mathcal{H}^1}(\Psi_{s,t,x}(t_0), \Psi'_{s,t,x}(t_0)). \quad (58)$$

- (vii) (*Quasi-rotation invariance*) If  $U \in SO(m)$  and  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WM}(I, E)$ , then

$$d_{S_\mu^1, I}(U\Psi_{s,t,x}, U\Psi'_{s,t,x}) \sim d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}). \quad (59)$$

Indeed, if we have metrics  $d_{S_\mu^1, I}$  on  $\text{WMC}(I, E)$  with the above properties, we can define metrics  $d_{S_\mu^1, I}$  on  $\text{WM}(I, E)$  by the usual Hausdorff distance construction

$$d_{S_\mu^1, I}(\phi, \phi') := \max\left(\sup_{\Psi_{s,t,x} \in \text{WMC}(\phi, I)} \inf_{\Psi' \in \text{WMC}(\phi', I)} d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'), \sup_{\Psi' \in \text{WMC}(\phi', I)} \inf_{\Psi \in \text{WMC}(\phi, I)} d_{S_\mu^1, I}(\Psi, \Psi')\right);$$

the verification that this construction allows us to deduce Theorem 4.1 from Theorem 5.11 is routine but somewhat tedious and is omitted.

It remains to establish Theorem 5.11. This is the purpose of the remaining sections of the paper.

## 6. PARABOLIC REGULARITY

The equations of motion (45), (46) show that  $\Psi_{s,t,x}$  obeys a parabolic equation in the  $s$  variable. Parabolic regularity theory then suggests that these fields should become increasingly smooth in  $s$  as  $s \rightarrow \infty$ . The

purpose of this section is to present an abstract formalisation of this principle, which we will use twice, once to control initial data, and again to control the solution itself. In order to establish the stability estimate (58), we will also need to adapt this regularity theory to differences between two solutions to the same parabolic equation.

Recall that  $\text{Schwartz}(I \times \mathbf{R}^2)$  is the space of smooth functions  $\phi : I \times \mathbf{R}^2 \rightarrow \mathbf{C}$  such that all derivatives are rapidly decreasing in space (with the usual topology). For each  $s > 0$ , recall that the integer  $k(s)$  is defined by (12).

**Definition 6.1** (Algebra family). Let  $I$  be a compact interval. An *algebra family* is a collection of continuous seminorms  $S_k(I \times \mathbf{R}^2)$  on  $\text{Schwartz}(I \times \mathbf{R}^2)$  that obeys the product estimate that obeys the following estimates:

- The product estimate

$$\|\phi^{(1)}\phi^{(2)}\|_{S_{\max(k_1, k_2)}(I \times \mathbf{R}^2)} \lesssim \|\phi^{(1)}\|_{S_{k_1}(I \times \mathbf{R}^2)} \|\phi^{(2)}\|_{S_{k_2}(I \times \mathbf{R}^2)} \quad (60)$$

for all  $k_1, k_2 \in \mathbf{Z}$  and  $\phi^{(1)}, \phi^{(2)} \in \text{Schwartz}(I \times \mathbf{R}^2)$ ;

- The parabolic regularity estimate

$$\|\partial_x^2 \phi(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim s^{-1} \|\phi(s/2)\|_{S_k(I \times \mathbf{R}^2)} + \sup_{s/2 \leq s' \leq s} \|(\partial_s - \Delta)\phi(s')\|_{S_k(I \times \mathbf{R}^2)} \quad (61)$$

for all smooth  $\phi : \mathbf{R}^+ \rightarrow \text{Schwartz}(I \times \mathbf{R}^2)$  and  $s > 0$ , as well as the variant

$$\|\partial_x^2 \phi(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim \|\partial_x^2 \phi(s'')\|_{S_k(I \times \mathbf{R}^2)} + \sup_{s'' \leq s' \leq s} \|(\partial_s - \Delta)\phi(s')\|_{S_k(I \times \mathbf{R}^2)} \quad (62)$$

for all smooth  $\phi : \mathbf{R}^+ \rightarrow \text{Schwartz}(I \times \mathbf{R}^2)$  and  $s/2 \leq s'' \leq s$ ;

- The additional parabolic regularity estimate

$$\|\partial_x^j e^{(s-s')\Delta} \phi\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_j (s'/s)^{\delta_1/10} (s-s')^{-j/2} \|\phi\|_{S_{k(s')}(I \times \mathbf{R}^2)} \quad (63)$$

for all smooth  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ ,  $s > s' \geq 0$ , and  $j \geq 0$ ;

- The comparability estimate

$$\|\phi\|_{S_{k_1}(I \times \mathbf{R}^2)} \lesssim \chi_{k_1=k_2}^{-\delta_1} \|\phi\|_{S_{k_2}(I \times \mathbf{R}^2)} \quad (64)$$

for all  $k_1, k_2 \in \mathbf{Z}$  and  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ .

- The uniform bound

$$\|\phi\|_{L_{t,x}^\infty(I \times \mathbf{R}^2)} \lesssim \|\phi\|_{S_k(I \times \mathbf{R}^2)} \quad (65)$$

for all  $k \in \mathbf{Z}$  and  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ .

*Example 6.2.* For a fixed time  $t_0$ , the norms

$$\|f\|_{S_k(I \times \mathbf{R}^2)} := \sup_{k' \in \mathbf{Z}} \chi_{k=k'}^{\delta_1} \sum_{j=0}^{10} 2^{-k'j} \|\partial_x^{j+1} P_{k'} f(t_0)\|_{L_x^2(\mathbf{R}^2)} + 2^{-k'j} \|\partial_x^j P_{k'} f(t_0)\|_{L_x^\infty(\mathbf{R}^2)}$$

form an algebra family.

**Theorem 6.3** (Abstract parabolic regularity). *Let  $I$  be an interval,  $S_k(I \times \mathbf{R}^2)$  be an algebra family. Let  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$  be differentiated fields, and let  $c, \delta c$  be frequency envelopes, with  $c$  having energy at most  $E$  for some  $E > 0$ . Suppose we have the bounds*

$$\|\partial_x^j \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+2)/2} \quad (66)$$

and

$$\|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c(s) s^{-(j+2)/2} \quad (67)$$

for all  $0 \leq j \leq 10$  and  $s > 0$ , where  $k(s)$  is defined in (12). Suppose also that we have the qualitative decay properties

$$\lim_{s \rightarrow \infty} s^{(j+1)/2} \|\partial_x^j \Psi_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} = 0 \quad (68)$$

for all  $j > 0$ . Then we have

$$\|\partial_x^j \Psi_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} c(s) s^{-(j+1)/2} \quad (69)$$

$$\|\partial_x^j A_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} c(s)^2 s^{-(j+1)/2} \quad (70)$$

$$\|\partial_x^j \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} c(s) s^{-(j+2)/2} \quad (71)$$

$$\|\partial_x^j \delta \Psi_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} \delta c(s) s^{-(j+1)/2} \quad (72)$$

$$\|\partial_x^j \delta A_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} c(s) \delta c(s) s^{-(j+1)/2} \quad (73)$$

$$\|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_{E,j} \delta c(s) s^{-(j+2)/2} \quad (74)$$

for all  $j \geq 0$  and  $s > 0$ . One can also obtain similar bounds with  $\partial_x^2$  replaced by  $\partial_s$  as one wishes (e.g. one could replace  $\partial_x^j$  by  $\partial_x^{j-4} \partial_s^2$  if  $j \geq 4$ ).

*Remark 6.4.* We have stated these bounds for arbitrary non-negative  $j$ , but in practice we will only need finitely many of these bounds (e.g. it would suffice to establish these bounds just in the range  $0 \leq j \leq 100$ ). Similarly for the remainder of the estimates in this section. Thus, the reader may safely ignore the dependence of  $j$  in the implied constants. For applications to the subsequent paper [27], it is worth remarking that the wave map equation  $D^\alpha \psi_\alpha = 0$  is not actually used at all in the argument here. We also make the technical remark that the connection  $A_x$  need not be Schwartz (see [25, Remark 3.14]), but in that case one can interpret the above bounds in the completion of the Schwartz space with respect to the  $S_k$  norm.

The rest of this section is devoted to the proof of Theorem 6.3. Let the notation and hypotheses be as in that theorem. We allow implied constants to depend on  $E$ , thus from (11) we have

$$\sup_{s>0} c(s) + \int_0^\infty c(s)^2 \frac{ds}{s} \lesssim 1. \quad (75)$$

Note that as the semi-norm  $S_k(I \times \mathbf{R}^2)$  is continuous in the Schwartz( $I \times \mathbf{R}^2$ ) topology, we can establish Minkowski's inequality

$$\left\| \int_{s_1}^{s_2} \psi(s) ds \right\|_{S_k(I \times \mathbf{R}^2)} \leq \int_{s_1}^{s_2} \|\psi(s)\|_{S_k(I \times \mathbf{R}^2)} ds \quad (76)$$

whenever the integral is absolutely convergent in the Schwartz( $I \times \mathbf{R}^2$ ) topology.

We begin with a bound on  $\Psi_x^*$  and  $\delta\Psi_x$ .

**Lemma 6.5.** *We have*

$$\|\partial_x^j \Psi_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+1)/2} \quad (77)$$

and

$$\|\partial_x^j \delta\Psi_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c(s) s^{-(j+1)/2} \quad (78)$$

for all  $0 \leq j \leq 9$  and  $s > 0$ .

*Proof.* We begin with (77). Let us write

$$f(s) := \sum_{j=0}^9 s^{(j+1)/2} \|\partial_x^j \Psi_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \quad (79)$$

for  $s > 0$ . To show (77), it suffices to show that  $f(s) \lesssim c(s)$  for all  $s > 0$ .

From (50), the Leibniz rule (4), and Minkowski's inequality (76) we have

$$\begin{aligned} f(s) &\lesssim \int_s^\infty \sum_{j=0}^9 s^{(j+1)/2} \|\partial_x^{j+1} \psi_s^*(s')\|_{S_{k(s)}(I \times \mathbf{R}^2)} \\ &\quad + \sum_{j_1, j_2 \geq 0: j_1 + j_2 \leq 9} s^{(j_1 + j_2 + 1)/2} \|\partial_x^{j_1} \Psi_x^*(s') \partial_x^{j_2} \psi_s^*(s')\|_{S_{k(s)}(I \times \mathbf{R}^2)} ds'. \end{aligned}$$

Using (66), (64), (60) we thus have

$$\begin{aligned} f(s) &\lesssim \int_s^\infty \sum_{j=0}^9 s^{(j+1)/2} c(s') (s')^{-(j+3)/2} (s'/s)^{\delta_1/2} \\ &\quad + \sum_{j_1, j_2 \geq 0: j_1 + j_2 \leq 9} s^{(j_1 + j_2 + 1)/2} (s')^{-(j_1 + 1)/2} f(s') c(s') (s')^{-(j_2 + 2)/2} (s'/s)^{-\delta_1/2} ds'. \end{aligned}$$

From (10), (75) we conclude

$$f(s) \lesssim c(s) + \int_s^\infty f(s') c(s') (s'/s)^{-\frac{3}{2} + \frac{\delta_1}{2} + \delta_0} \frac{ds'}{s'}.$$

From (68) we also see that  $f(s) \rightarrow 0$  as  $s \rightarrow \infty$ . By Lemma 2.15 and (10) we conclude that  $f(s) \lesssim c(s)$  as required.

Now we establish (78). We write

$$\delta f(s) := \sum_{j=0}^9 s^{(j+1)/2} \|\partial_x^j \delta \Psi_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)}$$

for  $s > 0$ . It suffices to show that  $f(s) \lesssim \delta c(s)$  for all  $s \gtrsim 0$ .

From (50) and (5) we have

$$\delta \Psi_x(s) = \int_s^\infty \mathcal{O}(\partial_x \delta \psi_s(s')) + \mathcal{O}((\delta \psi_s(s')) \Psi_x^*(s')) + \mathcal{O}(\psi_s^*(s') \delta \Psi_x(s')) ds'. \quad (80)$$

Using the Leibniz rule (4) and Minkowski's inequality (76), (66), (64), (77), and (60) we thus have

$$\begin{aligned} \delta f(s) &\lesssim \int_s^\infty \sum_{j=0}^9 s^{(j+1)/2} \delta c(s') (s')^{-(j+3)/2} (s'/s)^{\delta_1/2} \\ &\quad + \sum_{j_1, j_2 \geq 0: j_1 + j_2 \leq 9} s^{(j_1 + j_2 + 1)/2} (s')^{-(j_1 + 1)/2} c(s') \delta c(s') (s')^{-(j_2 + 2)/2} (s'/s)^{-\delta_1/2} ds' \\ &\quad + \sum_{j_1, j_2 \geq 0: j_1 + j_2 \leq 9} s^{(j_1 + j_2 + 1)/2} (s')^{-(j_1 + 1)/2} \delta f(s') c(s') (s')^{-(j_2 + 2)/2} (s'/s)^{-\delta_1/2} ds'. \end{aligned}$$

Using (10), (75) this simplifies to

$$\delta f(s) \lesssim \delta c(s) + \int_s^\infty \delta f(s') c(s') (s'/s)^{-\frac{3}{2} + \frac{\delta_1}{2} + \delta_0} \frac{ds'}{s'}.$$

By (68) we also have  $\delta f(s) \rightarrow 0$  as  $s \rightarrow \infty$ . The claim then follows from Lemma 2.15 as before.  $\square$

*Remark 6.6.* Note that every term that occurs when bounding  $\delta \Psi_x$  contains exactly one factor containing a  $\delta$ ; this is ultimately because the Leibniz rule (5) is “linear in  $\delta$ ”. This is why the final estimate is linear in  $\delta c$ . The same phenomenon will hold throughout the rest of this section.

We will need to compensate for the above loss of derivative (and for derivative losses in later parts of the argument). Here our main tools will be the parabolic equations (46), (45) and (61). We first recover one derivative for  $\psi_x$ :

**Lemma 6.7.** *We have*

$$\|\partial_x^j \psi_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+1)/2} \quad (81)$$

and

$$\|\partial_x^j \delta \psi_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c(s) s^{-(j+1)/2} \quad (82)$$

for all  $0 \leq j \leq 10$  and  $s > 0$ .

*Proof.* From Lemma 6.5 we only need to verify this when  $j = 10$ . We begin with (81). From (46) we have the schematic equation

$$\partial_s \psi_x^* = \Delta \psi_x^* + \mathcal{O}(\Psi_x^* \partial_x^* \Psi_x^*) + \mathcal{O}((\Psi_x^*)^3). \quad (83)$$

Applying the Leibniz rule (4) and (61) we conclude that

$$\begin{aligned} \|\partial_x^{10} \psi_x^*(s)\|_{S_k(I \times \mathbf{R}^2)} &\lesssim s^{-1} \|\partial_x^8 \psi_x^*(s/2)\|_{S_k(I \times \mathbf{R}^2)} \\ &\quad + \sup_{s/2 \leq s' \leq s} \sum_{j_1, j_2 \geq 0: j_1 + j_2 = 9} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \Psi_x^*\|_{S_k(I \times \mathbf{R}^2)} \\ &\quad + \sum_{j_1, j_2, j_3 \geq 0: j_1 + j_2 + j_3 = 8} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \Psi_x^* \partial_x^{j_3} \Psi_x^*\|_{S_k(I \times \mathbf{R}^2)}. \end{aligned}$$

Applying (60) and Lemma 6.5 we conclude

$$\|\partial_x^{10} \psi_x^*(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(10+1)/2} + \sup_{s/2 \leq s' \leq s} c(s')^2 s^{-(10+1)/2} + c(s')^3 s^{-(10+1)/2}$$

and the claim (81) follows from (10), (75).

Now we prove (82). From (46), (5) we have the schematic equation

$$\partial_s \delta \psi_x = \Delta \delta \psi_x + \mathcal{O}((\delta \Psi_x) \partial_x \Psi_x^*) + \mathcal{O}(\Psi_x^* \partial_x \delta \Psi_x) + \mathcal{O}((\Psi_x^*)^2 \delta \Psi_x).$$

Applying the Leibniz rule (4) and (61) we conclude that

$$\begin{aligned} \|\partial_x^{10} \delta \psi_x(s)\|_{S_k(I \times \mathbf{R}^2)} &\lesssim s^{-1} \|\partial_x^8 \delta \psi_x(s/2)\|_{S_k(I \times \mathbf{R}^2)} \\ &\quad + \sup_{s/2 \leq s' \leq s} \sum_{j_1, j_2 \geq 0: j_1 + j_2 = 9} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \delta \Psi_x\|_{S_k(I \times \mathbf{R}^2)} \\ &\quad + \sum_{j_1, j_2, j_3 \geq 0: j_1 + j_2 + j_3 = 8} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \Psi_x^* \partial_x^{j_3} \delta \Psi_x\|_{S_k(I \times \mathbf{R}^2)}. \end{aligned}$$

Applying (60), (81), Lemma 6.5, (10), and (75) we obtain the claim.  $\square$

Then we recover a derivative for  $A_x$  (gaining a useful factor of  $c(s)$  in the process):

**Lemma 6.8.** *We have*

$$\|\partial_x^j A_x^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s)^2 s^{-(j+1)/2} \quad (84)$$

and

$$\|\partial_x^j \delta A_x(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) \delta c(s) s^{-(j+1)/2} \quad (85)$$

for all  $0 \leq j \leq 10$  and  $s > 0$ .

*Proof.* We first prove (84). From (43), the Leibniz rule (4), and Minkowski's inequality (76) we have

$$\|\partial_x^j A_x^*\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \int_s^\infty \sum_{j_1, j_2 \geq 0: j_1 + j_2 = j} \|\partial_x^{j_1} \psi_s^*(s') \partial_x^{j_2} \psi_s^*(s')\|_{S_{k(s)}(I \times \mathbf{R}^2)} ds'.$$



Applying (60), Lemma 6.7, and (66) we thus have

$$\|\partial_x^j A_x^*\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \int_s^\infty c(s')^2 (s'/s)^{\delta_1/2} (s')^{-(j+3)/2} ds'.$$

Applying (10), (75), we obtain the claim (84).

Now we prove (85). From (43), (5), the Leibniz rule (4), and Minkowski's inequality (76) we have

$$\begin{aligned} \|\partial_x^j \delta A_x\|_{S_{k(s)}(I \times \mathbf{R}^2)} &\lesssim \int_s^\infty \sum_{j_1, j_2 \geq 0: j_1 + j_2 = j} \|\partial_x^{j_1} \psi_s^*(s') \partial_x^{j_2} \delta \psi_x(s')\|_{S_{k(s)}(I \times \mathbf{R}^2)} \\ &\quad + \|\partial_x^{j_1} (\delta \psi_s^*)(s') \partial_x^{j_2} \psi_x^*(s')\|_{S_{k(s)}(I \times \mathbf{R}^2)} ds'. \end{aligned}$$

Arguing as in the proof of (84) we obtain the claim.  $\square$

Now we recover a derivative for  $\psi_s$ :

**Lemma 6.9.** *We have*

$$\|\partial_x^j \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+2)/2} \quad (86)$$

and

$$\|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c(s) s^{-(j+2)/2} \quad (87)$$

for all  $0 \leq j \leq 11$  and  $s > 0$ .

*Proof.* From Proposition 6.5 we only need to verify this when  $j = 11$ . We begin with (86). From (45) we have the schematic equation

$$\partial_s \psi_s^* = \Delta \psi_s^* + \mathcal{O}(\Psi_x^* \partial_x \psi_s^*) + \mathcal{O}(\psi_s^* \partial_x \Psi_x^*) + \mathcal{O}((\Psi_x^*)^2 \psi_s^*). \quad (88)$$

Applying the Leibniz rule (4) and (61) we conclude that

$$\begin{aligned} \|\partial_x^{11} \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} &\lesssim s^{-1} \|\partial_x^9 \psi_s^*(s/2)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \\ &\quad + \sup_{s/2 \leq s' \leq s} \sum_{j_1, j_2 \geq 0: j_1 + j_2 = 10} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \psi_s\|_{S_{k(s)}(I \times \mathbf{R}^2)} \\ &\quad + \sum_{j_1, j_2, j_3 \geq 0: j_1 + j_2 = 9} \|\partial_x^{j_1} \Psi_x^* \partial_x^{j_2} \Psi_x^* \partial_x^{j_3} \psi_s^*\|_{S_{k(s)}(I \times \mathbf{R}^2)}. \end{aligned}$$

Applying (60), Lemma 6.7, Lemma 6.8, and (66) we thus have

$$\|\partial_x^{10} \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(11+2)/2} + \sup_{s/2 \leq s' \leq s} c(s')^2 s^{-(11+2)/2} + c(s')^3 s^{-(11+2)/2}.$$

Applying (10), (75) we obtain the claim (86).

Now we prove (87). From (45) and (5) we have the schematic equation

$$\begin{aligned} \partial_s \delta \psi_s &= \Delta \delta \psi_s + \mathcal{O}((\delta \Psi_x) \partial_x \psi_s^*) + \mathcal{O}(\Psi_x^* \partial_x \delta \psi_s) \\ &\quad + \mathcal{O}((\delta \psi_s) \partial_x \Psi_x^*) + \mathcal{O}(\psi_s^* \partial_x \delta \Psi_x^*) \\ &\quad + \mathcal{O}(\Psi_x^* (\delta \Psi_x) \psi_s^*) + \mathcal{O}((\Psi_x^*)^2 \delta \psi_s^*). \end{aligned} \quad (89)$$

As before, observe that every term on the right contains one  $\delta$  expression. The claim now follows by repeating the arguments used to prove (86).  $\square$

Note that the conclusion of this lemma is just like (66), but with one additional derivative of regularity on  $\psi_s$ . We can iterate these arguments indefinitely to conclude the claim of Theorem 6.3 with  $\partial_x$ . The analogous claims involving  $\partial_s$  can then be proven by further applications of estimates such as (88); we omit the (rather tedious) details.

## 7. INITIAL DATA ESTIMATES

In [25], various parabolic regularity estimates on the differentiated fields  $\Psi_{s,t,x} \in \text{WMC}(I, E)$  in the caloric gauge were established, which were either qualitative (with constants depending on  $\Psi_{s,t,x}$ ) or quantitative (with constants depending only on the energy  $E(\Psi_{s,t,x})$ ). In this section we develop the key estimates from that theory that we need. We begin with some qualitative estimates.

**Proposition 7.1** (Qualitative estimates). *If  $I$  is a compact interval and  $\Psi_{s,t,x} \in \text{WMC}(I, E)$  for some  $E > 0$ , then*

$$\|\partial_x^j \psi_{t,x}(s, t)\|_{L_x^1(\mathbf{R}^2)} \lesssim_{\phi, j} \langle s \rangle^{-j/2} \quad (90)$$

and

$$\|\partial_x^j \partial_{t,x} \psi_s(s, t)\|_{L_x^1(\mathbf{R}^2)} \lesssim_{\Psi_{s,t,x}, j} \langle s \rangle^{-(j+2)/2} \quad (91)$$

for all  $t \in I$ ,  $j \geq 0$ , and  $s \geq 0$ .

*Proof.* In [25, Theorem 3.12] and [25, Theorem 3.16], it was shown that  $\Psi_{s,t,x}$  was uniformly Schwartz in  $t, x$  for  $s$  in any bounded interval (here we use the compactness of  $I$ ), and so (91) already holds true when  $s$  is bounded.

It remains to handle the contribution of when  $s$  is large. Here we use the estimates

$$|\partial_t^j \partial_x^m \phi(s, t, x)| \lesssim_{m, j, \phi} \langle s \rangle^{-(m+1)/2}$$

for all  $j, m \geq 0$  just before [25, Equation (87)], which imply that

$$|\partial_t^j \partial_x^m \Psi_{t,x}(s, t, x)| \lesssim_{m, j, \phi} \langle s \rangle^{-(m+2)/2}. \quad (92)$$

Next, since  $\psi_{t,x}$  is Schwartz for bounded  $s$ , we have

$$\|\psi_{t,x}(s, t)\|_{L_x^1(\mathbf{R}^2)} \lesssim_{\phi} 1 \quad (93)$$

for all  $s = O(1)$ . Applying [24, Lemma 4.8] (noting that (46) that  $\psi_{t,x}$  obeys a covariant heat equation) we see that (93) in fact holds for all

$s$ . On the other hand, from (46) we have the heat equation

$$\partial_s \psi_{t,x} = \Delta \psi_{t,x} + \mathcal{O}(\partial_x(\Psi_x \psi_{t,x})) + \mathcal{O}((\partial_x \Psi_x) \psi_{t,x}) + \mathcal{O}(\Psi_x^2 \psi_{t,x}).$$

Repeatedly applying (92), Duhamel's formula and standard parabolic regularity estimates we then obtain (90) for all  $j \geq 0$ . Further application of (46) then gives

$$\|\partial_x^j \partial_s \psi_{t,x}(s, t)\|_{L_x^1(\mathbf{R}^2)} \lesssim_{\phi,j} \langle s \rangle^{-(j+2)/2};$$

using (52) we conclude (91).  $\square$

Now we turn to quantitative estimates.

**Theorem 7.2** (Initial data envelope bounds). *Let  $I$  be an interval,  $E > 0$ , and let  $t_0 \in I$ . Let  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$ . Then there exists frequency envelopes  $c_0, \delta c_0 : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  of energy  $O_E(1)$  and  $O_E(d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x}))^2$  respectively, such that*

$$\begin{aligned} & \|\partial_x^{j+1} \Psi_{t,x}^*(s, t)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \Psi_{t,x}^*(s, t)\|_{L_x^\infty(\mathbf{R}^2)} \\ & + \|\partial_x^j \partial_{t,x} \Psi_x^*(s, t)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \psi_s^*(s, t)\|_{L_x^2(\mathbf{R}^2)} \\ & + \|\partial_x^{j-1} \partial_{t,x} \psi_s^*(s, t)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j c_0(s) s^{-(j+1)/2} \end{aligned} \quad (94)$$

and

$$\begin{aligned} & \|\partial_x^{j+1} \delta \Psi_{t,x}(s, t)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \delta \Psi_{t,x}(s, t)\|_{L_x^\infty(\mathbf{R}^2)} \\ & + \|\partial_x^j \partial_{t,x} \delta \Psi_x(s, t)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \delta \psi_s(s, t)\|_{L_x^2(\mathbf{R}^2)} \\ & + \|\partial_x^{j-1} \partial_{t,x} \delta \psi_s(s, t)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j \delta c_0(s) s^{-(j+1)/2} \end{aligned} \quad (95)$$

for all  $j \geq 0$ , with the convention that we drop all terms involving  $\partial_x^{-1}$ .

The rest of the section is devoted to proving Theorem 7.2. We allow all implied constants to depend on  $E$ . We also work exclusively at time  $t$ , and omit explicit dependence on this time parameter.

We introduce the seminorm

$$\|f\|_{S_k(t)} := \sup_{k'} \chi_{k=k'}^{-\delta_1} \|P_{k'} \partial_{t,x} f(t)\|_{L_x^2(\mathbf{R}^2)}. \quad (96)$$

on  $\text{Schwartz}(I \times \mathbf{R}^2)$ . Note that this seminorm is continuous in the  $\text{Schwartz}(I \times \mathbf{R}^2)$  topology, and obeys (64).

We will define  $c_0$  by the formula

$$c_0(s) := \sup_{j=0}^{10} \sup_{s' > 0} \min((s'/s)^{\delta_0}, (s/s')^{\delta_0}) (s')^{(j+2)/2} \|\partial_x^j \psi_s^*(s', t)\|_{S_{k(s)}(t)}.$$

**Lemma 7.3.**  $c_0$  is a frequency envelope of energy  $O(1)$ . In particular we have the crude estimate

$$c_0(s) = O(1). \quad (97)$$

*Proof.* The property (10) is clear from construction, so it suffices to show that

$$\int_0^\infty c_0(s)^2 \frac{ds}{s} \lesssim 1.$$

If we set  $f_k(s) := \sum_{j=0}^{10} s^{(j+2)/2} \|\partial_x^j \psi_s^*(s, t)\|_{S_k(t)}$ , we see from the multiplicative form of Young's inequality (and (64)) that

$$\int_0^\infty c_0(s)^2 \frac{ds}{s} \lesssim \sum_k \sup_{s \sim 2^{-2k}} f_k(s)^2$$

and so by Poincaré's inequality, it suffices to show that

$$\int_0^\infty f_{k(s)}(s)^2 \frac{ds}{s} \lesssim 1. \quad (98)$$

and

$$\int_0^\infty (s f'_{k(s)}(s))^2 \frac{ds}{s} \lesssim 1. \quad (99)$$

From (54) we have

$$\psi_s^* = \mathcal{O}(\partial_x \psi_x^*) + \mathcal{O}((\Psi_x^*)^2)$$

and hence

$$\partial_{t,x} \partial_x^j \psi_s^* = \mathcal{O}(\partial_x^{j+1} \partial_{t,x} \psi_x^*) + \sum_{j_1+j_2=j} \mathcal{O}((\partial_x^{j_1} \Psi_x^*)(\partial_x^{j_2} \partial_{t,x} \Psi_x^*))$$

for  $0 \leq j \leq 10$ . From (53) we thus have

$$\partial_{t,x} \partial_x^j \psi_s^* = \mathcal{O}(\partial_x^{j+2} \psi_{t,x}^*) + \sum_{j_1+j_2=j} \mathcal{O}((\partial_x^{j_1} \Psi_x^*)(\partial_x^{j_2} \partial_{t,x} \Psi_x^*)) + \mathcal{O}((\partial_x^{j_1} \Psi_x^*)(\partial_x^{j_2} \partial_x \Psi_{t,x}^*)).$$

From (96), Littlewood-Paley theory and Bernstein's inequality we thus have

$$\begin{aligned} \|\partial_x^j \psi_s^*(s)\|_{S_k(I \times \mathbf{R}^2)} &\lesssim 2^k \|\partial_x^{j+1} \psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} + 2^{-k} \|\partial_x^{j+3} \psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \sum_{j_1+j_2=j, j+2} 2^{(j+1-j_1-j-1)k} \|\partial_x^{j_1} \Psi_x^*(s)\|_{L_x^2(\mathbf{R}^2)} \\ &\quad \times (\|\partial_x^{j_2} \partial_x \Psi_{t,x}^*(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j_2} \partial_{t,x} \Psi_x^*(s)\|_{L_x^\infty(\mathbf{R}^2)}). \end{aligned}$$

From [25, Lemma 7.2] we have

$$\|\partial_x^j \Psi_{t,x}^*\|_{L_x^\infty(\mathbf{R}^2)} \lesssim_j s^{-(j+1)/2}$$

for all  $j \geq 0$ ; applying (53) we conclude that

$$\|\partial_x^{j+1} \Psi_{t,x}^*\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \Psi_x^*\|_{L_x^\infty(\mathbf{R}^2)} \lesssim_j s^{-(j+1)/2}.$$

Inserting this into the previous estimates, we conclude that

$$f_{k(s)}(s) \lesssim \sum_{j=0}^{13} s^{j/2} \|\partial_x^j \Psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)};$$

a similar (but more involved) computation (using (51), (52), (54) to handle the additional  $s$  derivative) gives

$$s|f'_{k(s)}(s)| \lesssim \sum_{j=0}^{15} s^{j/2} \|\partial_x^j \Psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)}.$$

The claims (98), (99) then follow from [25, Corollary 4.6] and [25, Lemma 4.8].  $\square$

We similarly define

$$\delta c_0(s) := \sum_{j=0}^{10} \sup_{s' > 0} \min((s'/s)^{\delta_0}, (s/s')^{\delta_0}) (s')^{(j+2)/2} \|\partial_x^j \delta \psi_s(s', t)\|_{S_{k(s)}(t)},$$

then  $\delta c_0$  is also a frequency envelope of energy  $O(1)$ . We will obtain a better estimate on the energy of  $\delta c_0$  later.

We now plan to apply Theorem 6.3. From Littlewood-Paley theory and Bernstein's inequality we see that the  $S_{k(s)}$  norm obeys (60). Now we establish (61). We use Duhamel's formula (8) to write

$$\partial_x^2 \phi(s) = \partial_x^2 e^{s\Delta/2} \phi(s/2) + \int_{s/2}^s \partial_x^2 e^{(s-s')\Delta} (\partial_s - \Delta) \phi(s') ds'.$$

Direct calculation shows that the convolution kernel of  $\partial_x^2 e^{s\Delta/2}$  has an  $L^1$  norm of  $O(s^{-1})$ , and so the contribution of the first term is acceptable by Minkowski's inequality and the translation invariance of  $S_k(t)$ . So it suffices to show that

$$\left\| \int_{s/2}^s \partial_x^2 e^{(s-s')\Delta} (\partial_s - \Delta) \phi(s') ds' \right\|_{S_k(I \times \mathbf{R}^2)} \lesssim \sup_{s/2 \leq s' \leq s} \|(\partial_s - \Delta) \phi(s')\|_{S_k(I \times \mathbf{R}^2)}.$$

A direct application of Minkowski's inequality will give a logarithmic divergence, but we can take advantage of the Besov-space structure of  $S_k(t)$  to evade this. Indeed, from (96) (and the fact that Littlewood-Paley operators commute with partial derivatives) it suffices to show that

$$\left\| \int_{s/2}^s \partial_x^2 e^{(s-s')\Delta} (\partial_s - \Delta) P_{k'} \nabla_{t,x} \phi(s') ds' \right\|_{L_x^2(\mathbf{R}^2)} \lesssim \sup_{s/2 \leq s' \leq s} \|(\partial_s - \Delta) P_{k'} \nabla_{t,x} \phi(s')\|_{L_x^2(\mathbf{R}^2)}$$

for each  $k'$ . We may freely insert a projection  $P_{k'-5 < \cdot < k'+5}$  in front of the integrand. A calculation using the Fourier transform shows that the convolution kernel of  $P_{k'-5 < \cdot < k'+5} \partial_x^2 e^{(s-s')\Delta}$  has an  $L^1$  norm

of  $O(2^{2k'} \langle 2^{-2k'}(s-s') \rangle^{-100})$ , and so the claim (61) follows from translation invariance and Minkowski's inequality. Finally, from Proposition 7.1 and Bernstein's inequality we obtain the vanishing property (68).

We now have all the hypotheses of Theorem 6.3 verified; invoking this theorem, we obtain the estimates (69)-(74). From Littlewood-Paley theory and Bernstein's inequality we have

$$\|f(t)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_{t,x}f(t)\|_{L_x^2(\mathbf{R}^2)} \lesssim \|f\|_{S_k(t)}$$

for any  $k$ , and thus

$$\begin{aligned} \|\partial_x^j \Psi_x^*(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \Psi_x^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c_0(s) s^{-(j+1)/2} \\ \|\partial_x^j A_x^*(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} A_x^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c_0(s)^2 s^{-(j+1)/2} \\ \|\partial_x^j \psi_s^*(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \psi_s^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c_0(s) s^{-(j+2)/2} \\ \|\partial_x^j \delta \Psi_x(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \delta \Psi_x(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j \delta c_0(s) s^{-(j+1)/2} \\ \|\partial_x^j \delta A_x(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \delta A_x(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c_0(s) \delta c_0(s) s^{-(j+1)/2} \\ \|\partial_x^j \delta \psi_s(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^j \partial_{t,x} \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j \delta c_0(s) s^{-(j+2)/2} \end{aligned}$$

for all  $j \geq 0$  and  $s > 0$ . The estimates (94), (95) now follow from these estimates (and (51)-(54), the Leibniz rule, Hölder's inequality, and (97) as necessary).

It remains to establish the energy bound on  $\delta c_0$ .

**Lemma 7.4** (Energy bound). *We have*

$$\int_0^\infty (\delta c_0(s))^2 \frac{ds}{s} \lesssim d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2.$$

*Proof.* Write  $\mu_0 := d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})$  and  $\mu := (\int_0^\infty (\delta c_0(s))^2 \frac{ds}{s})^{1/2}$ . We may assume that  $\mu_0 \lesssim \mu$  as the claim is trivial otherwise. It will suffice to establish a bound of the form

$$\mu^2 \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon \tag{100}$$

for some  $0 < \varepsilon < 2$ .

By repeating the arguments used to prove Lemma 7.3, we have

$$\mu^2 \lesssim \int_0^\infty \delta f_{k(s)}(s)^2 \frac{ds}{s} + \int_0^\infty \delta (s f'_{k(s)}(s))^2 \frac{ds}{s}$$

where

$$\delta f_k(s) := \sum_{j=0}^{10} s^{(j+2)/2} \|\partial_x^j \delta \psi_s(s, t)\|_{S_k(t)}.$$

Continuing the arguments of Lemma 7.3 (adapted to differences, of course), we have

$$\delta f_{k(s)}(s) + s|\delta f'_{k(s)}(s)| \lesssim \sum_{j=0}^{15} s^{j/2} \|\partial_x^j \delta \Psi_{t,x}(s)\|_{L_x^2(\mathbf{R}^2)}$$

and thus

$$\mu^2 \lesssim \sum_{j=0}^{15} \int_0^\infty s^{j-1} \|\partial_x^j \delta \Psi_{t,x}(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds.$$

On the other hand, by (56) we have

$$\int_0^\infty \|\delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim \mu_0^2 \quad (101)$$

and

$$\|\delta \psi_t(0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \mu_0. \quad (102)$$

Meanwhile, from (95) we have

$$\|\partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j s^{-j/2} \delta c_0(s)$$

and

$$\|\partial_x^j \delta \psi_s(s)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim_j s^{-(j+1)/2} \delta c_0(s)$$

for all  $j \geq 0$ . Applying repeated  $s$  derivatives using the differenced versions of (45), (46), (51) we conclude

$$\|\partial_s^k \partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_{j,k} s^{-j/2-k} \delta c_0(s)$$

for all  $j, k > 0$ , and thus

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim_{j,k} \mu^2. \quad (103)$$

Interpolating with (101), we conclude that

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim_{j,k,\varepsilon} \mu_0^{2-\varepsilon} \mu^\varepsilon$$

for all  $j, k \geq 0$  and  $\varepsilon > 0$ . If we write

$$f(s) := \sum_{j=0}^{100} s^{(j+1)/2} \|\partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)}$$

we thus have

$$\int_0^\infty f(s)^2 \frac{ds}{s} + \int_0^\infty (s f'(s))^2 \frac{ds}{s} \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon$$

for all  $\varepsilon > 0$ . Arguing as in Lemma 7.3, we may thus find a frequency envelope  $\delta c'$  of energy  $O_\varepsilon(\mu_0^{2-\varepsilon} \mu^\varepsilon)$  for any  $\varepsilon > 0$ , such that  $f(s) \lesssim \delta c'(s)$  for all  $s > 0$ , thus

$$\|\partial_x^j \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim \delta c'(s) s^{-(j+1)/2}$$

for all  $s > 0$  and  $0 \leq j \leq 100$ . If we introduce the variant

$$\|f\|_{\tilde{S}_k(t)} := \sup_{k'} \chi_{k=k'}^{-\delta_1} \|P_{k'} \partial_x f(t)\|_{L_x^2(\mathbf{R}^2)}. \quad (104)$$

of (96), we conclude from Littlewood-Paley theory that

$$\|\partial_x^j \delta \psi_s(s)\|_{\tilde{S}_{k(s)}(t)} \lesssim \delta c'(s) s^{-(j+2)/2}$$

for all  $s > 0$  and  $0 \leq j \leq 10$ . Also, from (94) we have

$$\|\partial_x^j \psi_s^*(s)\|_{\tilde{S}_{k(s)}(t)} \lesssim c_0(s) s^{-(j+2)/2}$$

for the same range of  $s, j$ . We can then apply Theorem 6.3 as before and conclude that (69)-(74) hold for this norm. In particular, since

$$\|f(t)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x f(t)\|_{L_x^2(\mathbf{R}^2)} \lesssim \|f\|_{\tilde{S}_k(t)},$$

we have

$$\|\partial_x^j \delta \Psi_x(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j+1} \delta \Psi_x(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j \delta c'(s) s^{-(j+1)/2} \quad (105)$$

$$\|\partial_x^j \delta A_x(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j+1} \delta A_x(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j c_0(s) \delta c'(s) s^{-(j+1)/2} \quad (106)$$

$$\|\partial_x^j \delta \psi_s(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j+1} \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j \delta c'(s) s^{-(j+2)/2} \quad (107)$$

for all  $j \geq 0$  and  $s > 0$ .

These estimates will be adequate for controlling spatial derivatives. To control the time derivative, we must of course use (102). From (46) and (5) we know that  $\delta \psi_t$  obeys a heat equation,

$$\partial_s \delta \psi_t = \Delta \delta \psi_t + \mathcal{O}(A_x \partial_x \delta \psi_t) + \mathcal{O}(\partial_x A_x \delta \psi_t) + \mathcal{O}(\Psi_x^2 \delta \psi_t) + F \quad (108)$$

where the forcing term  $F$  has the form

$$F = \mathcal{O}((\delta A_x) \partial_x \psi_t) + \mathcal{O}(\partial_x \delta \Psi_x \psi_t) + \mathcal{O}(\Psi_x \delta \Psi_x \psi_t).$$

Integrating this heat equation against  $\delta \psi_t$  and using Hölder's inequality, we obtain the energy inequality

$$\begin{aligned} \partial_s \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2 &\leq -2 \|\partial_x \delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2 + O(\|A_x\|_{L_x^\infty(\mathbf{R}^2)} \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)} \|\partial_x \delta \psi_t\|_{L_x^2(\mathbf{R}^2)}) \\ &\quad + O((\|\partial_x A_x\|_{L_x^\infty(\mathbf{R}^2)} + \|\Psi_x\|_{L_x^\infty(\mathbf{R}^2)}^2) \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2) + O(\|F\|_{L_x^2(\mathbf{R}^2)} \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)}) \end{aligned}$$

for  $s > 0$ . From (94) we have

$$\begin{aligned} \|A_x\|_{L_x^\infty(\mathbf{R}^2)} &\lesssim c_0(s)^2 s^{-1/2} \\ \|\partial_x A_x\|_{L_x^\infty(\mathbf{R}^2)} + \|\Psi_x\|_{L_x^\infty(\mathbf{R}^2)}^2 &\lesssim c_0(s)^2 s^{-1} \end{aligned}$$

while from (94), (105)-(107), and Hölder's inequality we have

$$\|F\|_{L_x^2(\mathbf{R}^2)} \lesssim c_0(s) \delta c'(s) s^{-1}.$$

Using the elementary inequality  $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$  to split up some mixed terms, we obtain

$$\partial_s \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2 \leq -\|\partial_x \delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2 + O(c_0(s)^2 s^{-1} \|\delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2) + O((\delta c'(s))^2 s^{-1}).$$



Since  $c_0$  has energy  $O(1)$  and  $\delta c'(s)$  has energy  $O_\varepsilon(\mu_0^{2-\varepsilon}\mu^\varepsilon)$ , we conclude from Gronwall's inequality and (102) (discarding the negative  $\|\partial_x \delta \psi_t\|_{L_x^2(\mathbf{R}^2)}^2$  term) that

$$\|\delta \psi_t(s)\|_{L_x^2(\mathbf{R}^2)}^2 \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon$$

for all  $s > 0$  and  $\varepsilon > 0$ . Reinstating the discarded term, we then conclude that

$$\int_0^\infty \|\partial_x \delta \psi_t(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon.$$

From this, (108), Cauchy-Schwarz, and the preceding bounds we conclude that

$$\int_0^\infty \|(\partial_s - \Delta) \delta \psi_t\|_{L_x^2(\mathbf{R}^2)} ds \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon$$

and hence by Lemma 2.10 we have

$$\int_0^\infty \|\delta \psi_t(s)\|_{L_x^\infty(\mathbf{R}^2)}^2 ds \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon.$$

On the other hand, by repeating the arguments used to establish (103), we have

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^{j+1} \delta \psi_t(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim_{j,k} \mu^2$$

and

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^j \delta \psi_t(s)\|_{L_x^\infty(\mathbf{R}^2)}^2 ds \lesssim_{j,k} \mu^2$$

for all  $j, k \geq 0$ . Interpolating, we conclude that

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^{j+1} \delta \psi_t(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim_{j,k,\varepsilon} \mu_0^{2-\varepsilon} \mu^\varepsilon$$

and

$$\int_0^\infty s^{j+2k} \|\partial_s^k \partial_x^j \delta \psi_t(s)\|_{L_x^\infty(\mathbf{R}^2)}^2 ds \lesssim_{j,k,\varepsilon} \mu_0^{2-\varepsilon} \mu^\varepsilon.$$

for all  $j, k \geq 0$  and  $\varepsilon > 0$ . Arguing as with the construction of  $\delta c'$ , we may thus find a frequency envelope  $\delta c''$  of energy  $O_\varepsilon(\mu_0^{2-\varepsilon}\mu^\varepsilon)$  for any  $\varepsilon > 0$  such that

$$\|\partial_x^{j+1} \delta \psi_t(s)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \delta \psi_t(s)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \delta c''(s) s^{-(j+1)/2} \quad (109)$$

for all  $0 \leq j \leq 100$ . By increasing  $\delta c''$  if necessary we may assume that  $\delta c'' \geq \delta c'$ .

From (43) we have

$$\delta A_t(s) = \int_s^\infty \mathcal{O}(\delta \psi_t(s') \psi_s(s')) + \mathcal{O}(\psi_t(s') \delta \psi_s(s')) ds'$$

and so by (109), (107), the Leibniz rule, and the Minkowski and Hölder inequalities we have

$$\|\partial_x^{j+1} \delta A_t(s)\|_{L_x^2(\mathbf{R}^2)} + \|\partial_x^j \delta A_t(s)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \delta c''(s) s^{-(j+1)/2}$$

for all  $0 \leq j \leq 100$ . Combining this with (109), (105) we have

$$\|\partial_x^j \delta \Psi_{t,x}(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j+1} \delta \Psi_{t,x}(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim \delta c''(s) s^{-(j+1)/2}$$

and

$$\|\partial_x^j \delta \psi_s(s)\|_{L_x^\infty(\mathbf{R}^2)} + \|\partial_x^{j+1} \delta \psi_s(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim \delta c''(s) s^{-(j+2)/2}$$

for  $0 \leq j \leq 100$ . If we repeat the proof of Lemma 7.3, using the fact that  $\delta c''$  has energy  $O_\varepsilon(\mu_0^{2-\varepsilon} \mu^\varepsilon)$  we conclude that

$$\int_0^\infty \delta c_0(s)^2 \frac{ds}{s} \lesssim_\varepsilon \mu_0^{2-\varepsilon} \mu^\varepsilon.$$

Since the left-hand side is  $\mu^2$ , we obtain (100) as desired.  $\square$

The proof of Theorem 7.2 is now complete.

**7.5. Frequency localisation of  $\psi_s$ .** Heuristically,  $\psi_s(s)$  is concentrated at frequencies  $\sim s^{-1/2} \sim 2^{k(s)}$  with an energy of  $\sim s^{-1} c_0(s)$ . We make this heuristic more precise in the following technical lemma.

**Lemma 7.6.** *With the notation and assumptions as in Theorem 7.2, we have*

$$\sum_{k'} \chi_{k' \leq k(s)}^{-10} \chi_{k'=k(s)}^{-\delta_1} \|P_{k'} \partial_{t,x} \psi_s^*(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_E c_0(s) s^{-1} \quad (110)$$

and

$$\sum_{k'} \chi_{k' \leq k(s)}^{-10} \chi_{k'=k(s)}^{-\delta_1} \|P_{k'} \partial_{t,x} \delta \psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_E \delta c_0(s) s^{-1}. \quad (111)$$

for all  $s > 0$ .

*Proof.* Again we allow implied constants to depend on  $E$ . We use (53) to expand

$$\partial_{t,x} \psi_s^* = \mathcal{O}(\partial_x \partial_{t,x} \Psi_x^*) + \mathcal{O}(\Psi_x^* \partial_{t,x} \Psi_x^*).$$

Using (94) one already sees that

$$\|P_{k'} \mathcal{O}(\partial_x \partial_{t,x} \Psi_x^*)\|_{L_x^2(\mathbf{R}^2)} \lesssim \chi_{k' \geq k(s)} \chi_{k' \leq k(s)}^{20} c_0(s) s^{-1}$$

(say), and similarly that

$$\|\partial_x^{20} \mathcal{O}(\Psi_x^* \partial_{t,x} \Psi_x^*)\|_{L_x^2(\mathbf{R}^2)} \lesssim c_0(s) s^{-11}$$

and

$$\|\mathcal{O}(\Psi_x^* \partial_{t,x} \Psi_x^*)\|_{L_x^1(\mathbf{R}^2)} \lesssim c_0(s) s^{-1/2}$$

which by Bernstein's inequality gives

$$\|P_{k'}\mathcal{O}(\Psi_x^*\partial_{t,x}\Psi_x^*)\|_{L_x^2(\mathbf{R}^2)} \lesssim \chi_{k'\geq k(s)}\chi_{k'\leq k(s)}^{20}c_0(s)s^{-1}.$$

Putting this all together, we conclude (110). The claim (111) is proven by adapting the above argument to differences.  $\square$

We also need the following variant of the above lemma, controlling the lower frequency components of  $\psi_s$  and  $\psi_{t,x}$  more adequately:

**Lemma 7.7.** *With the notation and assumptions as in Theorem 7.2, we have*

$$\|P_k\nabla_{t,x}^i\psi_s^*(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_{E,j} c_0(s)s^{-(j+1)/2}\chi_{k\leq k(s)}^j\chi_{k=k(s)}^{0.1} \quad (112)$$

and

$$\|P_k\psi_{t,x}^*(0, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_E c_0(2^{-2k}), \quad (113)$$

and similarly

$$\|P_k\nabla_{t,x}^j\delta\psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_{E,j} \delta c_0(s)s^{-(j+1)/2}\chi_{k\leq k(s)}^j\chi_{k=k(s)}^{0.1} \quad (114)$$

and

$$\|P_k\delta\psi_{t,x}(0, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_E \delta c_0(2^{-2k}) \quad (115)$$

for all  $i = 0, 1$  and  $j \geq 0$ .

Note that these estimates are consistent with (56) and Lemma 7.4, but they give much more precise information about the frequency distribution of various components of the distance  $d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})$ .

*Proof.* We may rescale  $k = 0$ , and will omit the  $t_0$  subscript for brevity. We allow all implied constants to depend on  $E$ . From [25, Lemma 7.2] we record the basic estimates

$$\|\psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j s^{-j/2} \quad (116)$$

for all  $s > 0$ .

We first show (112). To simplify the notation we shall only handle the  $i = 0$  case, though the  $i = 1$  case follows (and is in fact somewhat easier, due to the increased decay in  $s$ ) using the same arguments. When  $s \geq 1$  this claim follows from Theorem 7.2, so assume  $s < 1$ ; our task is now to show that

$$\|P_0\psi_s^*(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim c_0(s)s^{-1/2+0.1}.$$

From (88) and the fundamental theorem of calculus we have

$$\begin{aligned} \|P_0\psi_s^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim \|P_0\psi_s^*(1)\|_{L_x^2(\mathbf{R}^2)} + \int_s^1 \|P_0\Delta\psi_s^*(s')\|_{L_x^2(\mathbf{R}^2)} + \|P_0(\Psi_x^*\partial_x\psi_s^*)(s')\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \|P_0(\psi_s^*\partial_x\Psi_x^*)(s')\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \|P_0((\Psi_x^*)^2\psi_s^*)(s')\|_{L_x^2(\mathbf{R}^2)} ds'. \end{aligned}$$

From (110) the first term is  $O(c_0(1))$ , which is acceptable by (10). By Theorem 7.2, the term  $\|P_0\Delta\psi_s^*(s')\|_{L_x^2(\mathbf{R}^2)}$  is  $O(c_0(s')(s')^{-1/2})$ , which is similarly acceptable by (10). For the term  $\|P_0(\Psi_x^*\partial_x\psi_s^*)(s')\|_{L_x^2(\mathbf{R}^2)}$ , we see from Theorem 7.2 and (116) that

$$\|\Psi_x^*\partial_x\psi_s^*(s')\|_{L_x^1(\mathbf{R}^2)} \lesssim c_0(s')(s')^{-1}$$

so by Bernstein's inequality

$$\|P_0(\Psi_x^*\partial_x\psi_s^*(s'))\|_{L_x^2(\mathbf{R}^2)} \lesssim c_0(s')(s')^{-1}$$

which is acceptable. Similar arguments dispose of the remaining terms in the integrand.

Now we show (113). From (46) we have

$$\partial_s\psi_{t,x}^* = \Delta\psi_{t,x}^* + \mathcal{O}(\Psi_{t,x}^*\partial_{t,x}^*\Psi_{t,x}^*) + \mathcal{O}((\Psi_{t,x}^*)^3).$$

so the fundamental theorem of calculus we have

$$\begin{aligned} \|P_0\psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim \|P_0\psi_{t,x}^*(1)\|_{L_x^2(\mathbf{R}^2)} + \int_s^1 \|P_0\Delta\psi_{t,x}^*(s')\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \|P_0(\Psi_{t,x}^*\partial_{t,x}^*\Psi_{t,x}^*)(s')\|_{L_x^2(\mathbf{R}^2)} + \|P_0((\Psi_{t,x}^*)^3)(s')\|_{L_x^2(\mathbf{R}^2)} ds'. \end{aligned}$$

From Theorem 7.2, the first term is  $O(c_0(1))$ , and the first integrand is  $O(c_0(s'))$ , which are both acceptable by (10). Repeating the previous Bernstein and interpolation arguments, one sees that all the remaining integrands are  $O(c_0(s')(s')^{-1.1})$ , which is also acceptable.

The remaining two estimates are obtained by adapting the above scheme to differences; we omit the details.  $\square$

## 8. SPACETIME FUNCTION SPACES

The metrics  $d_{S_\mu^1, I}$  needed for Theorem 4.1 will be constructed by applying certain spacetime function space norms (mostly from [20]) to the field  $\psi_s$ . Fortunately, we do not need to know the explicit construction of these spaces from [20] (which are rather complicated), but instead just need a certain abstract list of properties to be satisfied by these norms. In this section we record the properties we will need for these spaces (cf. [20, Theorem 3]). More precisely, in Section 10 we will establish

**Theorem 8.1** (Function space norms). *For every interval  $I$ , every integer  $k$ , and every  $\mu > 0$ , there exist translation-invariant norms  $S_k(I \times \mathbf{R}^2)$ ,  $S_{\mu,k}(I \times \mathbf{R}^2)$ ,  $N_k(I \times \mathbf{R}^2)$  on  $\text{Schwartz}(I \times \mathbf{R}^2)$ , with the following properties for all integers  $k, k_1, k_2, k_3$  and  $\phi, \phi^{(1)}, \phi^{(2)}, \phi^{(3)}, F \in \text{Schwartz}(I \times \mathbf{R}^2)$ :*

- (Continuity and monotonicity) If  $I = [t_-, t_+]$ , then  $\|\phi\|_{S_{\mu,k}([a,b])}$  is a continuous function of  $a, b$  for  $t_- \leq a < b \leq t_+$ , and is decreasing in  $a$  and increasing in  $b$ .
- ( $S_k$  and  $S_{\mu,k}$  are comparable) We have

$$\|\phi\|_{S_k(I \times \mathbf{R}^2)} \lesssim \|\phi\|_{S_{\mu,k}(I \times \mathbf{R}^2)} \lesssim \mu^{-1} \|\phi\|_{S_k(I \times \mathbf{R}^2)}. \quad (117)$$

- (Vanishing) If  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$  and  $t_0 \in I$ , then there exists an interval  $J \subset I$  containing  $t_0$  such that

$$\|\phi\|_{S_{\mu,k}(I \times \mathbf{R}^2)} \lesssim \sum_{k'} \chi_{k=k'}^{-\delta_1} \|\partial_{t,x} P_{k'} \phi(t_0)\|_{L_x^2(\mathbf{R}^2)}. \quad (118)$$

- (First product estimate) We have

$$\|\phi^{(1)} \phi^{(2)}\|_{S_{\max(k_1, k_2)}(I \times \mathbf{R}^2)} \lesssim \|\phi^{(1)}\|_{S_{k_1}(I \times \mathbf{R}^2)} \|\phi^{(2)}\|_{S_{k_2}(I \times \mathbf{R}^2)}. \quad (119)$$

- ( $N$  contains  $L_t^1 L_x^2$ ) If  $F \in \text{Schwartz}(I \times \mathbf{R}^2)$  has Fourier support in the region  $\{\xi : |\xi| \sim 2^k\}$ , then

$$\|F\|_{N_k(I \times \mathbf{R}^2)} \lesssim \|F\|_{L_t^1 L_x^2(I \times \mathbf{R}^2)}. \quad (120)$$

- (Adjacent  $N_k$  or  $S_k$  are equivalent) If  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ , then

$$\|\phi\|_{S_{k_1}(I \times \mathbf{R}^2)} \lesssim \chi_{k_1=k_2}^{-\delta_1} \|\phi\|_{S_{k_2}(I \times \mathbf{R}^2)} \quad (121)$$

and

$$\|\phi\|_{S_{\mu,k_1}(I \times \mathbf{R}^2)} \lesssim \chi_{k_1=k_2}^{-\delta_1} \|\phi\|_{S_{\mu,k_2}(I \times \mathbf{R}^2)}. \quad (122)$$

Similarly, if  $F \in \text{Schwartz}(I \times \mathbf{R}^2)$  and  $k_1 = k_2 + O(1)$ , then

$$\|F\|_{N_{k_1}(I \times \mathbf{R}^2)} \sim \|F\|_{N_{k_2}(I \times \mathbf{R}^2)}. \quad (123)$$

- (Energy estimate) If  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$  has Fourier transform supported in the region  $\{|\xi| \sim 2^k\}$ , and  $t_0 \in I$ , then

$$\|\phi\|_{S_k(I \times \mathbf{R}^2)} \lesssim \|\phi[t_0]\|_{\dot{H}^1(\mathbf{R}^2) \times L^2(\mathbf{R}^2)} + \|\square \phi\|_{N_k(I \times \mathbf{R}^2)}. \quad (124)$$

- (Parabolic regularity estimate) If  $\phi : \mathbf{R}^+ \rightarrow \text{Schwartz}(I \times \mathbf{R}^2)$  is smooth and  $s > 0$ , then

$$\|\partial_x^2 \phi(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim s^{-1} \|\phi(s/2)\|_{S_k(I \times \mathbf{R}^2)} + \sup_{s/2 \leq s' \leq s} \|(\partial_s - \Delta) \phi(s')\|_{S_k(I \times \mathbf{R}^2)} \quad (125)$$

and similarly for  $S_{\mu,k}(I \times \mathbf{R}^2)$  or  $N_k(I \times \mathbf{R}^2)$ .

- (Second product estimate) We have

$$\|P_k(\phi F)\|_{N_k(I \times \mathbf{R}^2)} \lesssim \chi_{k \geq k_2}^{\delta_2} \chi_{k=\max(k_1, k_2)}^{\delta_2} \|\phi\|_{S_{k_1}(I \times \mathbf{R}^2)} \|F\|_{N_{k_2}(I \times \mathbf{R}^2)}. \quad (126)$$

- (Improved trilinear estimate) We have

$$\begin{aligned} \|P_k(\phi^{(1)} \partial_\alpha \phi^{(2)} \partial^\alpha \phi^{(3)})\|_{N_k(I \times \mathbf{R}^2)} &\lesssim \mu^{1-\varepsilon} \chi_{k=\max(k_1, k_2, k_3)}^{\varepsilon \delta_1} \chi_{k_1 \leq \min(k_2, k_3)}^{\varepsilon \delta_1} \\ &\quad \times \|\phi^{(1)}\|_{S_{\mu_1, k_1}(I \times \mathbf{R}^2)} \|\phi^{(2)}\|_{S_{\mu_2, k_2}(I \times \mathbf{R}^2)} \|\phi^{(3)}\|_{S_{\mu_3, k_3}(I \times \mathbf{R}^2)} \end{aligned} \quad (127)$$

and

$$\begin{aligned} \|P_k(\phi^{(1)}\partial_\alpha\phi^{(2)}\partial^\alpha\phi^{(3)})\|_{N_k(I\times\mathbf{R}^2)} &\lesssim \mu^{2-2\varepsilon}\chi_{k=\max(k_1,k_2,k_3)}^{\varepsilon\delta_1}\chi_{k_1\leq\min(k_2,k_3)}^{\varepsilon\delta_1} \\ &\quad \|\phi^{(1)}\|_{S_{\mu,k_1}(I\times\mathbf{R}^2)}\|\phi^{(2)}\|_{S_{\mu,k_2}(I\times\mathbf{R}^2)}\|\phi^{(3)}\|_{S_{\mu,k_3}(I\times\mathbf{R}^2)} \end{aligned} \quad (128)$$

for every  $0 \leq \varepsilon \leq 1$  and if  $k_1 \geq \min(k_2, k_3) - O(1)$ , whenever two of the  $\mu_1, \mu_2, \mu_3$  are equal to  $\mu$  and the third is equal to 1 (with the convention that  $S_{1,k} = S_k$ ).

- (Strichartz estimates) If  $\phi$  has Fourier support in the region  $\{\xi : |\xi| \lesssim 2^k\}$ , then we have

$$\sup_{t \in I} \|\phi[t]\|_{H^1(\mathbf{R}^2) \times L_x^2(\mathbf{R}^2)} \lesssim \|\phi\|_{S_k(I \times \mathbf{R}^2)} \quad (129)$$

and

$$\|\partial_{t,x}\phi\|_{L_t^5 L_x^\infty(I \times \mathbf{R}^2)} \lesssim 2^{4k/5} \mu \|\phi\|_{S_{\mu,k}(I \times \mathbf{R}^2)} \quad (130)$$

and

$$\|\partial_{t,x}\phi\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim 2^{(1-1/q)k} \|\phi\|_{S_k(I \times \mathbf{R}^2)}. \quad (131)$$

for  $5 \leq q \leq \infty$ .

*Remark 8.2.* One could form the metric completion of the Schwartz space  $\text{Schwartz}(I \times \mathbf{R}^2)$  under the above norms to obtain Banach spaces instead of mere normed vector spaces, but we will not need to do so here as our analysis will remain purely in the Schwartz category (note that Theorem 5.11 deals exclusively with classical solutions).

*Remark 8.3.* We make the trivial but very useful remark that all the above estimates for scalar-valued functions automatically extend to vector or tensor-valued functions of bounded dimension (with a slight degradation in the implied constants). We will frequently use this remark in the rest of the paper (especially when dealing with expressions in schematic form) without future comment.

*Remark 8.4.* The trilinear estimates (127), (128) are variants of the estimate

$$\begin{aligned} \|P_k(\phi^{(1)}\partial_\alpha\phi^{(2)}\partial^\alpha\phi^{(3)})\|_{N_k(I\times\mathbf{R}^2)} &\lesssim \chi_{k=\max(k_1,k_2,k_3)}^{\delta_1}\chi_{k_1\leq\min(k_2,k_3)}^{\delta_1} \\ &\quad \|\phi^{(1)}\|_{S_{k_1}(I\times\mathbf{R}^2)}\|\phi^{(2)}\|_{S_{k_2}(I\times\mathbf{R}^2)}\|\phi^{(3)}\|_{S_{k_3}(I\times\mathbf{R}^2)} \end{aligned}$$

that was established (with some difficulty) in [20, Section 18], and indeed we will use that estimate in the proof of (127), (128). The key improvement in (127), (128), which is crucial for the large data theory, is that we can gain an additional factor of the parameter  $\mu$ .

*Remark 8.5.* The  $L_t^5 L_x^\infty$  Strichartz estimate in (130) is a little short of the endpoint estimate  $L_t^4 L_x^\infty$ . We were not able to establish this estimate for our function spaces (the energy estimate with null frame atom forcing term was problematic); fortunately, as observed in [9],

[10], non-endpoint Strichartz estimates such as  $L_t^5 L_x^\infty$  estimates are still available, and suffice for the purpose of estimating higher order terms. Fortunately, the Strichartz estimates are only needed for higher order terms anyway, and in fact any non-trivial  $L_t^q L_x^\infty$  Strichartz estimate (where by “non-trivial” we mean that  $q$  is finite) would suffice for our purposes.

We close this section with some basic corollaries of Theorem 8.1. Firstly, we make the technical observation that the  $S_k(I \times \mathbf{R}^2)$ ,  $S_{\mu,k}(I \times \mathbf{R}^2)$ ,  $N_k(I \times \mathbf{R}^2)$  norms are continuous with respect to the Schwartz( $I \times \mathbf{R}^2$ ) topology (which, in particular, will allow us to use Minkowski’s inequality (76) in those norms when the integrands are uniformly Schwartz). For the  $N_k(I \times \mathbf{R}^2)$  norm this follows from (120). For the  $S_k(I \times \mathbf{R}^2)$  norm, we use Littlewood-Paley decomposition, the triangle inequality, (124), and (121), (120) to obtain the useful estimate

$$\|\phi\|_{S_k(I \times \mathbf{R}^2)} \lesssim \sum_{k'} \chi_{k=k'}^{-\delta_1} [\|P_{k'} \phi[t_0]\|_{\dot{H}^1(\mathbf{R}^2) \times L^2(\mathbf{R}^2)} + \|\square P_{k'} \phi\|_{L_t^1 L_x^2(I \times \mathbf{R}^2)}]. \quad (132)$$

It is not difficult to show that the right-hand side goes to zero when  $\phi$  goes to zero in Schwartz( $I \times \mathbf{R}^2$ ), thus establishing continuity of the  $S_k(I \times \mathbf{R}^2)$  norm. The claim for  $S_{\mu,k}(I \times \mathbf{R}^2)$  then follows from (117).

Next, we observe the interpolation estimate

$$\|\partial_x^{j'} \phi\|_{S_k(I \times \mathbf{R}^2)} \lesssim_{j',j} \|\phi\|_{S_k(I \times \mathbf{R}^2)}^{1-j'/j} \|\partial_x^j \phi\|_{S_k(I \times \mathbf{R}^2)}^{j'/j} \quad (133)$$

for all  $0 < j' < j$  and  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ , and similarly for the  $S_{\mu,k}(I \times \mathbf{R}^2)$  and  $N_k$  norms. Indeed, given any frequency parameter  $k_0$ , we can use Littlewood-Paley theory to write

$$\partial_x^{j'} \phi = \partial_x^{j'} P_{\leq k_0} \phi + \partial_x^{j'-j} P_{> k_0} \partial_x^j \phi,$$

where  $\partial_x^{j'-j}$  is a suitable Fourier multiplier of order  $j' - j$ . It is not hard to see that  $\partial_x^{j'} P_{\leq k_0}$  and  $\partial_x^{j'-j} P_{> k_0}$  are convolution operators whose kernel has  $L^1$  norm  $O_{j,j'}(2^{j'k_0})$  and  $O_{j,j'}(2^{(j'-j)k_0})$  respectively, so from Minkowski’s inequality (76), the triangle inequality, and the translation invariance of the  $S_k(I \times \mathbf{R}^2)$  norms we conclude

$$\|\partial_x^{j'} \phi\|_{S_k(I \times \mathbf{R}^2)} \lesssim_{j',j} 2^{j'k_0} \|\phi\|_{S_k(I \times \mathbf{R}^2)} + 2^{(j'-j)k_0} \|\partial_x^j \phi\|_{S_k(I \times \mathbf{R}^2)}$$

for any  $k_0$ . Optimising in  $k_0$  we obtain the claim.

## 9. PROOF OF THEOREM 5.11 ASSUMING THEOREM 8.1

In this (lengthy) section we assume Theorem 8.1 and use it to prove Theorem 5.11. Fix  $E > 0$ ; all constants are allowed to depend on  $E$ .

**9.1. Definition of the metrics.** For each integer  $k$  and  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$ , define the quantity

$$d_{\mu,k,I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) := \sum_{j=0}^{10} \sup_{2^{-2k-2} \leq s \leq 2^{-2k}} s^{1+\frac{j}{2}} \|\partial_x^j(\psi_s(s) - \psi'_s(s))\|_{S_{\mu,k}(I \times \mathbf{R}^2)} \quad (134)$$

and then define the metric  $d_{S_{\mu}^1, I}$  by the formula

$$d_{S_{\mu}^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) := \left( \sum_k d_{\mu,k,I}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2 \right)^{1/2} \quad (135)$$

To ensure that these are actually metrics, we have to check finiteness and non-degeneracy:

**Lemma 9.2** (Finiteness). *For any  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$ ,  $\tilde{d}_{S_{\mu}^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x})$  is finite.*

*Proof.* By the triangle inequality it suffices to show that the sequence

$$\sup_{2^{-2k-2} \leq s \leq 2^{-2k}} s^{1+\frac{j}{2}} \|\partial_x^j \psi_s(s)\|_{S_{\mu,k}(I \times \mathbf{R}^2)}$$

is square-summable in  $k$  for  $0 \leq j \leq 10$  and  $\Psi_{s,t,x} \in \text{WMC}(I, E)$ .

By (117) we may replace  $S_{\mu,k}$  by  $S_k$  here. Applying (132) we have

$$\begin{aligned} \|\partial_x^j \psi_s(s)\|_{S_k(I \times \mathbf{R}^2)} &\lesssim \sum_{k'} \chi_{k=k'}^{-\delta_1} [\|\partial_x^j \partial_{t,x} P_{k'} \psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \|\partial_x^j \square P_{k'} \psi_s(s)\|_{L_t^1 L_x^2(I \times \mathbf{R}^2)}] \end{aligned} \quad (136)$$

where  $t_0 \in I$  is arbitrary.

From Proposition 7.1 and Bernstein's inequality we have

$$\|\partial_x^j \partial_{t,x} P_{k'} \psi_s(s, t)\|_{L_x^2(\mathbf{R}^2)} \lesssim_{\phi,j,m} \min(2^{k'}, (\langle s \rangle 2^{2k'})^{-m}) \langle s \rangle^{-(j+2)/2}$$

for any  $m$ . This ensures that the contribution of the first term in (136) is acceptable.

As for the second term, we first recall from [25, Lemma 7.5] that

$$\|\partial_x^j \partial_s w(s, t)\|_{L_x^1(\mathbf{R}^2)} \lesssim_{j,\phi} s^{-(j+2)/2} \quad (137)$$

for  $s \gtrsim 1$ . From (48) we have

$$\square \psi_s = \mathcal{O}(\partial_s w) + \mathcal{O}(\partial_{t,x} \Psi_{t,x} \psi_s) + \mathcal{O}(\Psi_{t,x} \partial_{t,x} \psi_s) + \mathcal{O}(\Psi_{t,x}^2 \psi_s).$$

Using (137), Proposition 7.1, (92) we conclude that

$$\|\partial_x^j \square \psi_s\|_{L_x^1(\mathbf{R}^2)} \lesssim_{\phi,j} s^{-(j+2)/2}$$

for all  $j \geq 0$  and  $s \gtrsim 1$ , which is enough to show that the second term in (136) is acceptable.  $\square$



**Lemma 9.3** (Non-degeneracy). *If  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$  are such that  $\tilde{d}_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) = 0$ , then  $\Psi_{s,t,x} = \Psi'_{s,t,x}$ .*

*Proof.* This is clear from construction (note from (say) (130) that the only functions with vanishing  $S_{\mu,k}(I \times \mathbf{R}^2)$  norm are identically zero).  $\square$

**9.4. Easy verifications.** We now verify some of the easier components of Theorem 5.11. The monotonicity of  $\|\Psi_{s,t,x}\|_{S_\mu^1(I \times \mathbf{R}^2)}$  in  $I$  follows from the monotonicity of the  $S_{\mu,k}$  norms from Theorem 8.1. The continuity of  $\|\Psi_{s,t,x}\|_{S_\mu^1(I \times \mathbf{R}^2)}$  in  $I$  similarly follows from the continuity of the  $S_{\mu,k}$  norms and the dominated convergence theorem (using Lemma 9.2 to provide the domination). The quasi-isometry property (59) follows easily by breaking everything up into components and using the triangle inequality.

Now we show the vanishing property. Fix  $\Psi_{s,t,x}, \Psi'_{s,t,x}$ . By Theorem 7.2 we can find frequency envelopes  $c_0, \delta c_0$  with the stated properties.

By construction of  $d_{S_\mu^1, I_n}$ , it will suffice to show that

$$\lim_{n \rightarrow \infty} \sum_k d_{\mu,k,I}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2 \lesssim d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2.$$

From (134), (118) (and continuity of the  $S_{\mu,k,I}$  norm in the Schwartz topology), we know that

$$\lim_{n \rightarrow \infty} d_{\mu,k,I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) \lesssim \sum_{j=0}^{10} \sup_{2^{-2k-2} \leq s \leq 2^{-2k}} s^{1+\frac{j}{2}} \sum_{k'} \chi_{k=k'}^{-\delta_1} \|\partial_x^j \partial_{t,x} P_{k'}(\psi_s - \psi'_s)(s, t_0)\|_{L_x^2(\mathbf{R}^2)}.$$

Thus by the monotone convergence theorem, it suffices to show that

$$\sum_k \left[ \sup_{2^{-2k-2} \leq s \leq 2^{-2k}} s^{1+\frac{j}{2}} \sum_{k'} \chi_{k=k'}^{-\delta_1} \|\partial_x^j \partial_{t,x} P_{k'}(\psi_s - \psi'_s)(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \right]^2 \lesssim d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2.$$

But by (111), the expression in brackets is  $O(\delta c_0(2^{-k}))$ , and the claim follows from (10) and the fact that  $\delta c_0$  has energy  $O(d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2)$ .

**9.5. The stability priori estimate.** We now establish (58). Fix  $M, \mu, I, t_0, \Psi_{s,t,x}, \Psi'_{s,t,x}$ , where we assume  $\mu$  sufficiently small depending on  $M$ . We allow all implied constants to depend on  $M$ , thus

$$\|\Psi_{s,t,x}\|_{S_\mu^1(I \times \mathbf{R}^2)}, \|\Psi'_{s,t,x}\|_{S_\mu^1(I \times \mathbf{R}^2)} \lesssim 1.$$

Our task is to show that

$$d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x}) \lesssim_\mu d_{\mathcal{H}^1}(\phi_{s,t,x}(t_0), \phi'_{s,t,x}(t_0)). \quad (138)$$

By (135), we have

$$\sum_k d_{\mu,k,I}(\Psi_{s,t,x}, 0)^2, \sum_k d_{\mu,k,I}(\Psi'_{s,t,x}, 0)^2 \lesssim 1. \quad (139)$$

and our task is to show that

$$\sum_k d_{\mu,k,I}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2 \lesssim_\mu d_{\mathcal{H}^1}(\Psi_{s,t,x}(t_0), \Psi'_{s,t,x}(t_0))^2.$$

In view of (117), it suffices to show that

$$\sum_k d_k(\Psi_{s,t,x}, \Psi'_{s,t,x})^2 \lesssim d_{\mathcal{H}^1}(\Psi_{s,t,x}(t_0), \Psi'_{s,t,x}(t_0))^2 \quad (140)$$

where

$$d_k(\Psi_{s,t,x}, \Psi'_{s,t,x}) := \sum_{j=0}^{10} \sup_{2^{-2k-2} \leq s \leq 2^{-2k}} s^{1+\frac{j}{2}} \|\partial_x^j(\psi_s(s) - \psi'_s(s))\|_{S_k}.$$

Now define

$$c(s) := \sum_k [d_{\mu,k,I}(\Psi_{s,t,x}, 0) + d_{\mu,k,I}(\Psi'_{s,t,x}, 0)] \min((2^{2k}s)^{\delta_0}, (2^{2k}s)^{-\delta_0}) \quad (141)$$

and

$$\delta c(s) := \sum_k d_k(\Psi_{s,t,x}, \Psi'_{s,t,x}) \min((2^{2k}s)^{\delta_0}, (2^{2k}s)^{-\delta_0}). \quad (142)$$

Then  $c, \delta c$  are frequency envelopes, and by (139) and Young's inequality we see that  $c$  has energy  $O(1)$ . In particular from (9), (11) we have

$$\sup_s c(s) + \int_0^\infty c(s)^2 \frac{ds}{s} \lesssim 1. \quad (143)$$

From (142), we see that to prove (140), it will suffice to show that

$$\int_0^\infty \delta c(s)^2 \frac{ds}{s} \lesssim d_{\mathcal{H}^1}(\phi[t_0], \phi'[t_0])^2. \quad (144)$$

From (141), (142), (121), (139), we see that<sup>9</sup>

$$\|\partial_x^j \psi_s^*(s)\|_{S_{\mu,k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+2)/2} \quad (145)$$

and

$$\|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c(s) s^{-(j+2)/2} \quad (146)$$

---

<sup>9</sup>There will be many powers of  $s$  on the right-hand side of the estimates in this section, but one does not need to pay too much attention to the exponents here, as they are always equal to the exponent predicted by the dimensional analysis heuristics  $\psi_{t,x}, \partial_{t,x}, A_{t,x}, 2^k \sim s^{-1/2}$ ;  $\psi_s \sim s^{-1}$ ;  $L_t^q \sim s^{-1/(2q)}$ ;  $L_x^r \sim s^{-1/r}$ , with  $P_k, c(s), \mu$  being dimensionless.

for all  $0 \leq j \leq 10$  and  $s > 0$ , where  $k(s)$  was defined in (12). From (145), (117) we of course have

$$\|\partial_x^j \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-(j+2)/2} \quad (147)$$

for the same range of  $j, s$ .

The arguments used to prove Lemma 9.2 also establish the vanishing property (68). We can now invoke Theorem 6.3 to obtain the estimates (69)-(74) for all  $j \geq 0$ . We also have analogous bounds for  $S_{\mu,k}$ , for  $\psi_s^*$  at least:

**Corollary 9.6** (Infinite gain of regularity in  $S_{\mu,k}$ ).

$$\|\partial_x^j \psi_s^*(s)\|_{S_{\mu,k(s)}(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{-\varepsilon} c(s) s^{-(j+2)/2} \quad (148)$$

for all  $j \geq 0$ ,  $\varepsilon > 0$  and  $s > 0$ .

*Proof.* From (146) we have (148) for  $j \leq 10$  (with no loss of  $\mu^{-\varepsilon}$ ), while from (71) and (117) we have (148) for all  $j$  with a loss of  $\mu^{-O(1)}$ . Interpolating using (133) we conclude (148) for all  $j$  and all  $\varepsilon > 0$ .  $\square$

This, together with (145), leads to the following Strichartz estimates which will be useful for disposing of higher order terms in the nonlinearity.

**Lemma 9.7** (Strichartz estimates). *We have*

$$\|\partial_x^j \partial_{t,x} \Psi_x^*(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q} c(s) s^{-(j+2)/2+1/(2q)} \quad (149)$$

$$\|\partial_x^j \psi_s^*(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q} c(s) s^{-(j+2)/2+1/(2q)} \quad (150)$$

$$\|\partial_x^j \partial_{t,x} \psi_s^*(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q} c(s) s^{-(j+4)/2+1/(2q)} \quad (151)$$

$$\|\partial_x^j \Psi_{t,x}^*(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q} c(s) s^{-(j+1)/2+1/(2q)} \quad (152)$$

$$\|\partial_x^j A_{t,x}^*(s)\|_{L_t^r L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/2r} c(s) s^{-(j+1)/2+1/(2r)} \quad (153)$$

$$\|\partial_x^j \partial_{t,x} \Psi_x^*(s)\|_{L_t^\infty L_x^2(I \times \mathbf{R}^2)} \lesssim_j c(s) s^{-(j+1)/2} \quad (154)$$

$$\|\partial_x^j \psi_s^*(s)\|_{L_t^\infty L_x^2(I \times \mathbf{R}^2)} \lesssim_j c(s) s^{-(j+1)/2} \quad (155)$$

$$\|\partial_x^j \partial_{t,x} \psi_s^*(s)\|_{L_t^\infty L_x^2(I \times \mathbf{R}^2)} \lesssim_j c(s) s^{-(j+2)/2} \quad (156)$$

for all  $j \geq 0$ ,  $s > 0$ ,  $5 \leq q \leq \infty$ ,  $5/2 \leq r \leq \infty$ , and  $\varepsilon > 0$ . One also has analogues of all the above estimates in which  $\Psi^*$  is replaced by  $\delta\Psi$ ,  $c$  is replaced by  $\delta c$ , etc., and all powers of  $\mu$  are discarded. Thus for instance, the analogue of (149) is

$$\|\partial_x^j \partial_{t,x} \delta\Psi_x(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)} \lesssim_j \delta c(s) s^{-(j+2)/2+1/(2q)}. \quad (157)$$

*Proof.* We first establish the estimates for  $\Psi_{s,t,x}^*$ , and return to  $\delta\Psi_{s,t,x}$  later.

From Corollary 9.6, (121) we have

$$\|P_k \Psi_x^*(s)\|_{S_{\mu,k}(I \times \mathbf{R}^2)} \lesssim_{j',\varepsilon} \mu^{-\varepsilon} \chi_{k=k(s)}^{-\delta_1} c(s) s^{-1/2} \langle 2^k s^{-1/2} \rangle^{-j'}$$

for all  $k, j', \varepsilon$ , and similarly for the  $S_k(I \times \mathbf{R}^2)$  norm without the  $\mu^{-\varepsilon}$  loss (thanks to (69)). By (130), (131) we thus have

$$\|\partial_x^j \partial_{t,x} P_k \Psi_x^*\|_{L_t^5 L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,j'} \mu^{1-\varepsilon} 2^{kj} 2^{(1-1/5)k} c(s) s^{-1/2} \langle 2^k s^{-1/2} \rangle^{-j'}$$

and

$$\|\partial_x^j \partial_{t,x} P_k \Psi_x^*\|_{L_t^\infty L_x^\infty(I \times \mathbf{R}^2)} \lesssim_{j,j'} 2^{kj} 2^k c(s) s^{-1/2} \langle 2^k s^{-1/2} \rangle^{-j'}$$

for any  $j, j'$ . Interpolating these estimates, taking  $j' = j + 5$  (say) and summing in  $k$ , we conclude (149).

From (149), (38), the Leibniz rule (4), and Hölder's inequality we obtain (150), (151).

Now we turn to (152). If we set

$$f_{j,q}(s) := \sum_{j'=0}^j c(s)^{-1} s^{(j+1)/2-1/2q} \|\partial_x^j \Psi_{t,x}^*(s)\|_{L_t^q L_x^\infty(I \times \mathbf{R}^2)}$$

then from (49), Minkowski's inequality, the Leibniz rule, Hölder's inequality, (150), (151), and (10) we conclude that

$$f_{j,q}(s) \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q} + \int_{s'}^\infty (s'/s)^{-\sigma} f_{j,q}(s') c(s') ds' / s'$$

for some absolute constant  $\sigma > 0$  (independent of  $\delta_0$ ) and all  $\varepsilon > 0$ . Noting from Proposition 7.1, Bernstein's inequality, and Hölder's inequality that  $f_{j,q}(s) \rightarrow 0$  as  $s \rightarrow \infty$ . Applying Lemma 2.15 we conclude that  $f_{j,q}(s) \lesssim_{j,\varepsilon} \mu^{(5-\varepsilon)/q}$  for all  $s > 0$  and  $\varepsilon > 0$ , and (152) follows.

The estimate (153) follows from (43), the Leibniz rule, Minkowski's inequality, Hölder's inequality, (150), (152), and (10), (143).

The estimate (154) follows from (69) and (129). The estimates (155), (156) then follows from (154), (152), and (38).

The analogous estimates for  $\delta\Psi$  are obtained by applying (5) to all the equations of motion used above, and modifying the arguments appropriately (without trying to gain any factors of  $\mu$ ).  $\square$

Having obtained adequate control on  $\Psi_{s,t,x}^*, \Psi'_{s,t,x}$ , we now turn attention to the wave-tension fields  $w = D^\alpha \psi_\alpha$ ,  $w' = (D')^\alpha \psi'_\alpha$ ,  $\delta w = w' - w$ . We begin with the nonlinear forcing term

$$F := (\psi_\alpha \wedge \psi_i) D_i \psi^\alpha \quad (158)$$

appearing in (47), with  $F'$  (and hence  $F^*$  and  $\delta F$ ) defined accordingly. Here we can obtain estimates which (crucially) gains more than one power of  $\mu$ , as well as some decay at low frequencies.

**Lemma 9.8** (Forcing term estimate). *We have*

$$\|P_k F^*(s)\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{2-\varepsilon} c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2} \chi_{k \leq k(s)}^j \quad (159)$$

and

$$\|P_k \delta F(s)\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{1-\varepsilon} \kappa c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2} \chi_{k \leq k(s)}^j \quad (160)$$

for all  $j \geq 0$ ,  $s > 0$ , and integers  $k$ , and  $\varepsilon > 0$ .

*Proof.* We will need to extract the implicit null structure from the expression  $F$  by exploiting “dynamic separation” as in [10], [9]. In the context of the caloric gauge, dynamic separation entails using (44), (43) to rewrite the terms involving  $\alpha$  in (158) in terms of  $\psi_s$ , modulo higher order terms. Indeed, from (44) we can express  $F(s)$  as the sum of the cubic term

$$\int_s^\infty \int_s^\infty \mathcal{O}(\partial_\alpha \psi_s(s') \psi_x(s) \partial_x \partial^\alpha \psi_s(s'')) ds' ds'' \quad (161)$$

the quintic<sup>10</sup> terms

$$\int_s^\infty \int_s^\infty \mathcal{O}(\partial_\alpha \psi_s(s') \psi_x(s) A_x(s) \partial^\alpha \psi_s(s'')) ds' ds'' \quad (162)$$

$$\int_s^\infty \int_s^\infty \mathcal{O}(A_\alpha(s') \psi_s(s') \psi_x(s) \partial_x \partial^\alpha \psi_s(s'')) ds' ds'' \quad (163)$$

$$\int_s^\infty \int_s^\infty \mathcal{O}(\partial_\alpha \psi_s(s') \psi_x(s) \partial_x (A^\alpha(s'') \psi_s(s''))) ds' ds'' \quad (164)$$

the septic terms

$$\int_s^\infty \int_s^\infty \mathcal{O}(A_\alpha(s') \psi_s(s') \psi_x(s) A_x(s) \partial^\alpha \psi_s(s'')) ds' ds'' \quad (165)$$

$$\int_s^\infty \int_s^\infty \mathcal{O}(A_\alpha(s') \psi_s(s') \psi_x(s) \partial_x (A^\alpha(s'') \psi_s(s''))) ds' ds'' \quad (166)$$

$$\int_s^\infty \int_s^\infty \mathcal{O}(\partial_\alpha \psi_s(s') \psi_x(s) A_x(s) A^\alpha(s'') \psi_s(s'')) ds' ds'' \quad (167)$$

and the nonic term

$$\int_s^\infty \int_s^\infty \mathcal{O}(A_\alpha(s') \psi_s(s') \psi_x(s) A_x(s) A^\alpha(s'') \psi_s(s'')) ds' ds''. \quad (168)$$

Note in each of these expressions, the derivatives such as  $\partial_\alpha$  or  $\partial_x$  are falling on “low frequency” terms (terms arising from large values  $s', s''$  of the heat-temporal parameter, rather than from small values such as

<sup>10</sup>Here we count  $A$  as a quadratic term (cf. (43), (84); also note the range of exponents in (153) is twice as large as (149)-(151)).

$s$ ). This phenomenon, which also occurs in the integral expressions appearing later in this proof, will be crucial in ensuring that the integrals are convergent in the required function space norms.

We now turn to the proof of (159). It suffices to show that

$$\|P_k \partial_x^j F^*(s)\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{2-O(\varepsilon)} c(s) s^{-(j+2)/2} \chi_{k \geq k(s)}^{\delta_2}$$

for all  $j \geq 0$ ,  $s > 0$ ,  $k \in \mathbf{Z}$ ,  $\varepsilon > 0$ . We split  $F^*$  into terms of the form (161)-(168) (but with  $\Psi_{s,t,x}$  replaced by  $\Psi_{s,t,x}^*$ , etc.).

We first deal with the cubic term (161). By Minkowski's inequality, the contribution of this term is bounded by

$$\lesssim \int_s^\infty \int_s^\infty \|P_k \partial_x^j (\partial_\alpha \psi_s^*(s') \psi_x^*(s) \partial_x \partial^\alpha \psi_s^*(s''))\|_{N_k(I \times \mathbf{R}^2)} ds' ds''.$$

Applying the Leibniz rule (4), (128), (69), Corollary 9.6, and (10), (143), we can bound this by

$$\lesssim_{j,\varepsilon} \int_s^\infty \int_s^\infty s^{-j/2} \mu^{2-O(\varepsilon)} c(s) \chi_{k=k(s)}^{\varepsilon \delta_1} (\max(s', s'')/s)^{-\varepsilon \delta_1/2} (s')^{-1} s^{-1/2} (s'')^{-3/2} ds' ds'';$$

performing the integrals we see that these terms are acceptable. Note that the  $j > 0$  cases are no harder than the  $j = 0$  case (and in some terms there is even a slight gain); the reader may in fact wish to set  $j = 0$  when following the discussion below, as the higher  $j$  case never adds any substantial new difficulty.

The first quintic term (162) can be handled similarly. Indeed, from (69), (119), (143), and the Leibniz rule (4), we have

$$\|\partial_x^j (\psi_x(s) A_x(s))\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_j c(s) s^{-j/2} s^{-1}$$

for all  $j \geq 0$ , and then by repeating the previous arguments we can estimate the contribution of this case by

$$\lesssim_{j,\varepsilon} \int_s^\infty \int_s^\infty s^{-j/2} \mu^{2-O(\varepsilon)} c(s) \chi_{k=k(s)}^{\varepsilon \delta_1} (\max(s', s'')/s)^{-\varepsilon \delta_1/2} (s')^{-1} s^{-1} (s'')^{-1} ds' ds''$$

which is also (barely) acceptable.

We set aside the other quintic terms for now and look at the nonic term (168). By Minkowski's inequality, (120), and Bernstein's inequality, and discarding the null structure, the contribution of this term is bounded by

$$\lesssim 2^k \int_s^\infty \int_s^\infty \|P_k \partial_x^j (A_{t,x}^*(s') \psi_s^*(s') \psi_x^*(s) A_x^*(s) A_{t,x}^*(s'') \psi_s^*(s''))\|_{L_t^1 L_x^1(I \times \mathbf{R}^2)} ds' ds''.$$

We apply Leibniz's rule (4) and Hölder's inequality, estimating the “high frequency” terms  $\psi_x(s)$ ,  $A_x^*(s)$  in  $L_t^\infty L_x^2$ , the term  $A_{t,x}^*(s')$  (say)

in  $L_t^{5/2}L_x^\infty$ , and the other three terms  $A_{t,x}^*(s'')$ ,  $\psi_s(s')$ ,  $\psi_s(s'')$  in  $L_t^5L_x^\infty$ . Applying Lemma 9.7, (143) we can estimate the above expression by

$$\lesssim_{j,\varepsilon} 2^k \int_s^\infty \int_s^\infty s^{-j/2} \mu^{2-O(\varepsilon)} c(s)(s')^{-3/2+1/5} (s'')^{-3/2+3/10} ds' ds''.$$

(in fact we have several powers of  $\mu$  to spare). Performing the integrals we see that this term is acceptable.

The septic terms can be treated by the same methods as the nonic term, but with a more delicate numerology. To illustrate this, consider for instance (167). If we do not apply Bernstein's inequality, but otherwise repeat the above argument, we can control the contribution of this term by

$$\lesssim \int_s^\infty \int_s^\infty \|P_k \partial_x^j (\partial_{t,x}(s') \psi_s^*(s') \psi_x^*(s) A_x^*(s) A_{t,x}^*(s'') \psi_s^*(s''))\|_{L_t^1 L_x^2(I \times \mathbf{R}^2)} ds' ds''.$$

If we place  $\psi_x^*(s)$  in  $L_t^\infty L_x^2$ ,  $A_x^*(s)$  and  $A_{t,x}^*(s'')$  in  $L_t^{10/3} L_x^\infty$  (say), and  $\partial_{t,x}(s') \psi_s^*(s')$  and  $\psi_s^*(s'')$  in  $L_t^5 L_x^\infty$ , we obtain the integral

$$\lesssim_{j,\varepsilon} \int_s^\infty \int_s^\infty s^{-j/2} \mu^{2-O(\varepsilon)} c(s)(s')^{-3/2+1/10} s^{-1/2+3/20} (s'')^{-3/2+1/4} ds' ds''$$

(again, we have powers of  $\mu$  to spare). This integral is almost acceptable, except that we did not gain the factor of  $\chi_{k \geq k(s)}^{\delta_2}$ . For this, we must invoke Bernstein's inequality a little bit, moving  $L_t^1 L_x^2$  to  $L_t^1 L_x^{2-c}$  for some small  $c$ , and then moving some of the exponents in Hölder's inequality on, say,  $A_x^*(s)$ , to match (by interpolating the various estimates in Lemma 9.7); note that scale-invariance assures us that we end up with the right exponent of  $s$  in the end. We leave the details to the reader. The treatment of the terms (165), (166) are similar (with some permutations in the exponents) and are also left to the reader.

The quintic terms (163), (164) cannot be directly treated by Strichartz methods. (Even with the endpoint  $L_t^4 L_x^\infty$  Strichartz estimate, one would barely be able to place the nonlinearity in  $L_t^1 L_x^2$ , leaving no room for the Bernstein inequality to give the important factor of  $\chi_{k \geq k(s)}^{\delta_2}$ . This issue also comes up in [10].) To deal with this, we need to perform more dynamic separation on these terms. For instance, to deal with (163), we use (43), (44) to write

$$A_\alpha^*(s') = \int_{s'}^\infty \int_{s'''}^\infty \mathcal{O}(\psi_s^*(s''') \partial_\alpha \psi_s^*(s''')) + \mathcal{O}(\psi_s^*(s''') A_\alpha^*(s''') \psi_s^*(s''')) ds''' ds''.$$

This splits (163) into a quintic null form

$$\int_s^\infty \int_s^\infty \int_{s'}^\infty \int_{s'''}^\infty \mathcal{O}(\psi_s^*(s''') \partial_\alpha \psi_s^*(s''') \psi_s^*(s') \psi_x^*(s) \partial_x \partial^\alpha \psi_s^*(s'')) ds''' ds'' ds' ds''$$

and a septic expression

$$\int_s^\infty \int_s^\infty \int_{s'}^\infty \int_{s'''}^\infty \mathcal{O}(\psi_s^*(s''') A_\alpha^*(s''') \psi_s^*(s''') \psi_s^*(s') \psi_x^*(s) \partial_x \partial^\alpha \psi_s^*(s'')) ds'''' ds''' ds' ds''.$$

The septic expression can be dealt with by the same sort of Strichartz and Bernstein techniques as the previous septic expressions (165), (166), (167), while the quintic null form can be dealt with similarly to the quintic null form (162). Similar arguments let one handle (164). We omit the details as they are very similar to the expressions already dealt with, except for some minor permutations in the numerology (the key point again being that the derivatives are falling on low frequency terms rather than high frequency ones).

This ends our discussion of (159). To prove (160), one applies (5) to all the expressions above, thus replacing one of the  $\Psi^*$  factors appearing above with a  $\delta\Psi$  factor (for various values of “ $\Psi$ ”). One then repeats the above arguments, using the appropriate “ $\delta$  versions” of estimates such as (69) or Lemma 9.7, and using (127) instead of (128). Because the estimates on  $\delta$  terms do not have any gains in  $\mu$  (and in particular are estimated in  $S_k$  rather than  $S_{\mu,k}$ ), we only end up with a power of  $\mu^{1-O(\varepsilon)}$  rather than  $\mu^{2-O(\varepsilon)}$ . Instead of keeping a factor of  $c(s)$  in the final estimate, one instead retains a factor such as  $\delta c(s)$ ,  $\delta c(s')$ ,  $\delta c(s'')$ , etc. (depending on where the  $\delta$  term has fallen); however, by using (10) one can convert this back to  $\delta c(s)$  at a cost of an expression such as  $(s'/s)^{\delta_0}$ , which does not affect the convergence of any of the integrals (as they already have some room to spare in these exponents, typically of the order of  $1/10$  or so). We again leave the details to the reader, as they are somewhat tedious.  $\square$

We are now in a position to obtain control on the wave-tension field  $w = D^\alpha \psi_\alpha$ , that also gains the crucial  $\mu$  factor:

**Lemma 9.9** (Control of wave-tension field). *We have*

$$\|P_k w^*\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{2-\varepsilon} c(s) \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (169)$$

and

$$\|P_k \partial_s w^*\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{2-\varepsilon} c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (170)$$

for all  $j \geq 0$ ,  $s > 0$ ,  $\varepsilon > 0$ , and integers  $k$ . Similarly we have

$$\|P_k \delta w\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{1-\varepsilon} \delta c(s) \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (171)$$

and

$$\|P_k \partial_s \delta w\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \mu^{1-\varepsilon} \delta c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (172)$$

for the same range of  $j, s, \varepsilon, k$ .



*Proof.* We begin with the proof of (170). Fix  $j$ . Without loss of generality we may take  $j \geq 10$  (say).

From (47) we have the schematic heat equation

$$\partial_s w^* = \Delta w^* + \mathcal{O}(A_x^* \partial_x w^*) + \mathcal{O}((\partial_x A_x^*) w^*) + \mathcal{O}((\Psi_x^*)^2 w^*) + \mathcal{O}(F^*). \quad (173)$$

From Duhamel's formula (8) and (39) we conclude that

$$w^*(s) = \int_0^s e^{(s-s')\Delta} (\mathcal{O}(A_x^* \partial_x w^*) + \mathcal{O}((\partial_x A_x^*) w^*) + \mathcal{O}((\Psi_x^*)^2 w^*) + \mathcal{O}(F^*))(s') ds'. \quad (174)$$

Let

$$f_j(s) := \sup_k c(s)^{-1} \chi_{k \geq k(s)}^{-\delta_2/10} \chi_{k \leq k(s)}^{-j} \|P_k w^*\|_{N_k(I \times \mathbf{R}^2)},$$

thus

$$\|P_k w^*\|_{N_k(I \times \mathbf{R}^2)} \lesssim f_j(s) c(s) \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (175)$$

for all  $k$ . The claim (169) is then equivalent to showing that  $f_j(s) \lesssim_{j,\varepsilon} \mu^{2-\varepsilon}$  for all  $s > 0$ ,  $j \geq 0$ , and  $\varepsilon > 0$ .

Observe that the convolution kernel of  $P_k e^{(s-s')\Delta}$  has total mass  $O_j(\chi_{k \leq k(s-s')}^{100j})$  for any  $j$ . From (174) and Minkowski's inequality, we can thus bound  $f_j(s)$  by

$$\lesssim_j \sup_k c(s)^{-1} \chi_{k \geq k(s)}^{-\delta_2} \chi_{k \leq k(s)}^{-j} \int_0^s \chi_{k \leq k(s-s')}^{100j} \times \\ \|P_k(\mathcal{O}(A_x^* \partial_x w^*) + \mathcal{O}((\partial_x A_x^*) w^*) + \mathcal{O}((\Psi_x^*)^2 w^*) + \mathcal{O}(F^*))(s')\|_{N_k(I \times \mathbf{R}^2)} ds'.$$

For any  $s' > 0$ , we see from (69), (70), (119) that

$$\|\partial_x A_x^*(s')\|_{S_{k(s')}(I \times \mathbf{R}^2)}, \|(\Psi_x^*)^2(s')\|_{S_{k(s')}(I \times \mathbf{R}^2)} \lesssim c(s')^2 (s')^{-1}$$

and thus by (175), (126) and dyadic decomposition we have

$$\|P_k(\mathcal{O}((\partial_x \Psi_x^*) w^*) + \mathcal{O}((\Psi_x^*)^2 w^*))(s')\|_{N_k(I \times \mathbf{R}^2)} \lesssim_j (s'/s)^{\delta_2/2} c(s')^3 f_j(s') (s')^{-1} \chi_{k \geq k(s')}^{\delta_2/10} \chi_{k \leq k(s')}^j.$$

A similar argument also gives

$$\|P_k(\mathcal{O}(\Psi_x^* \partial_x w^*))\|_{N_k(I \times \mathbf{R}^2)} \lesssim_j c(s')^3 (s'/s)^{\delta_2/2} f_j(s') (s')^{-1} \chi_{k \geq k(s')}^{\delta_2/10} \chi_{k \leq k(s')}^{j-1}.$$

Combining these estimates with Lemma 9.8, we conclude that

$$f_j(s) \lesssim_{j,\varepsilon} \sup_k c(s)^{-1} \chi_{k \geq k(s)}^{-\delta_2/10} \chi_{k \leq k(s)}^{-j} \\ \int_0^s \chi_{k \leq k(s-s')}^{100j} [(s'/s)^{\delta_2/2} c(s')^3 f_j(s') \chi_{k \geq k(s')}^{\delta_2/10} \chi_{k \leq k(s')}^{j-1} \\ + \mu^{2-\varepsilon} c(s') \chi_{k \geq k(s')}^{\delta_2} \chi_{k \leq k(s')}^{100j}] \frac{ds'}{s'}$$

for all  $\varepsilon > 0$ .

A direct computation using (10) and the elementary bounds

$$\chi_{k \geq k(s)}^{-\delta_2/10} \leq \chi_{k \geq k(s')}^{-\delta_2/10}, \quad \chi_{k \leq k(s)}^{-j} \lesssim_j \chi_{k \leq k(s-s')}^{-j} \chi_{k \leq k(s')}^{-j}$$

shows that

$$c(s)^{-1} \chi_{k \geq k(s)}^{-\delta_2/10} \chi_{k \leq k(s)}^{-j} \int_0^s \chi_{k \leq k(s-s')}^{100j} c(s') \chi_{k \geq k(s')}^{\delta_2} \chi_{k \leq k(s')}^{100j} \frac{ds'}{s'} \lesssim_j 1.$$

Using this, (10), and the elementary bounds

$$\begin{aligned} \chi_{k \geq k(s)}^{-\delta_2/10} \chi_{k \geq k(s')}^{\delta_2/10} &\lesssim (s'/s)^{-\delta_2/20} \\ \chi_{k \leq k(s)}^{-j} &\lesssim \chi_{k \leq k(s-s')}^{-j} \chi_{k \leq k(s')}^{-j} \\ \chi_{k \leq k(s-s')} \chi_{k \leq k(s')}^{-1} &\lesssim ((s-s')/s)^{-1/2} \end{aligned}$$

then gives the integral inequality

$$f_j(s) \lesssim_{j,\varepsilon} \mu^{2-\varepsilon} + \int_0^s ((s-s')/s)^{-1/2} (s'/s)^{\delta_2/4} c(s')^2 f_j(s') \frac{ds'}{s'}$$

for all  $\varepsilon > 0$ . Also, from (120), Proposition 7.1, Bernstein's inequality, and Hölder's inequality we have  $f_j(s) \rightarrow 0$  as  $s \rightarrow 0$ . Applying Lemma 2.16 we have  $f_j(s) \lesssim_{j,\varepsilon} \mu^{2-\varepsilon}$ , and the claim (169) follows.

The claim (170) then follows by using (173) to write  $\partial_s w^*$  in terms of  $\partial_x^2 w^*$  (which can be controlled by (169)), together with several additional terms which were already estimated in the required manner in the first part of the proof.

The proof of (171) follows from (169) as in previous propositions. A little more specifically, we define

$$\delta f_j(s) := \sup_k \delta c(s)^{-1} \chi_{k \geq k(s)}^{-\delta_2/10} \chi_{k \leq k(s)}^{-j} \|P_k \delta w\|_{N_k(I \times \mathbf{R}^2)}$$

and repeat the above arguments (and using the bound (169) just established) to eventually obtain the integral inequality

$$\delta f_j(s) \lesssim_{j,\varepsilon} \mu^{1-\varepsilon} + \int_0^s ((s-s')/s)^{-1/2} (s'/s)^{\delta_2/4} c(s')^2 \delta f_j(s') \frac{ds'}{s'}$$

Using the continuity method as before we obtain  $\delta f_j(s) \lesssim_{j,\varepsilon} \mu^{1-\varepsilon}$  for all  $s > 0$ ,  $j \geq 0$ , and  $\varepsilon > 0$ , giving (171). The proof of (172) then follows by the differenced version of (173). We leave the details to the reader.  $\square$

Having earned the crucial gains in  $\mu$ , we now return to control  $\psi_s$ .

**Lemma 9.10** (Return to  $\psi_s$ ). *We have*

$$\|P_k \psi_s^*(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \|\partial_{t,x} P_k \psi_s^*(s, t_0)\|_{L_x^2(\mathbf{R}^2)} + \mu^{2-\varepsilon} c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (176)$$

and

$$\|P_k \delta \psi_s(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim_{j,\varepsilon} \|\partial_{t,x} P_k \delta \psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} + \mu^{1-\varepsilon} \delta c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^j \quad (177)$$

for all  $j \geq 0$ ,  $\varepsilon > 0$ ,  $s > 0$ , and integers  $k$ .

*Remark 9.11.* The estimate (176) will not be directly needed for the stability estimate (26), but is useful for the persistence of regularity estimate (27) in the next section.

*Proof.* We begin by proving (176). Let us fix  $j, k, s, \varepsilon$ , and suppress dependence of implied constants on  $j, \varepsilon$ . From (48) we have the schematic wave equation

$$\square \psi_s^*(s) = \mathcal{O}(G_1) + \mathcal{O}(G_2) + \mathcal{O}(G_3) + \mathcal{O}(G_4) + \mathcal{O}(G_5)$$

where

$$\begin{aligned} G_1 &:= A_\alpha^*(s) \partial^\alpha \psi_s^*(s) \\ G_2 &:= (\partial^\alpha A_\alpha^*(s)) \psi_s^*(s) \\ G_3 &:= (A^*)^\alpha(s) A_\alpha^*(s) \psi_s^*(s) \\ G_4 &:= (\psi^*)^\alpha(s) \psi_\alpha^*(s) \psi_s^*(s) \\ G_5 &:= \partial_s w^*(s). \end{aligned}$$

Applying (124), we have

$$\|P_k \psi_s^*(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim \|\partial_{t,x} P_k \psi_s^*(s, t_0)\|_{L_x^2(\mathbf{R}^2)} + \sup_{1 \leq i \leq 5} \|P_k G_i\|_{N_{k'}(I \times \mathbf{R}^2)}.$$

It thus suffices to show that

$$\|\partial_x^j P_k G_i\|_{N_k(I \times \mathbf{R}^2)} \lesssim \mu^{2-\varepsilon} c(s) s^{-(j+2)/2} \chi_{k \geq k(s)}^{\delta_2/10} \quad (178)$$

for  $i = 1, 2, 3, 4, 5$ .

The term  $G_5$  can be treated by Lemma 9.9. To handle the other terms  $G_1, G_2, G_3, G_4$ , it will be necessary to use dynamic separation to make the null structure in these expressions more explicit, creating cubic and quintic expressions with explicit null structure, together with sextic and higher order terms which can be handled by Strichartz estimates. Unfortunately, this process creates a large number<sup>11</sup> of terms to estimate, and so it will be convenient to introduce some notation to partially unify the treatment of these terms.

---

<sup>11</sup>With the decomposition we use, there are 28 (!) terms, though symmetry allows one to identify a few of these terms together. One could cut down the number of terms somewhat by establishing a certain quintilinear estimate from the function spaces  $S_k, S_{*,k}$  to  $N_k$ , as remarked earlier, thus relieving the need to deal with septilinear terms, but this does not seem to result in a net gain in simplicity for the paper and we have not pursued this approach.

Given any integers  $J \geq 0$  and  $\tilde{J}, J' \geq 1$  and formal symbols  $D, D' \in \{\partial, A^*\}$ , we let  $G_{\tilde{J}, J, J'}^{D, D'}$  denote the schematic quantity

$$G_{\tilde{J}, J, J'}^{D, D'} := \int_* \psi_s^*(\tilde{s}_1) \dots \psi_s^*(\tilde{s}_{\tilde{J}}) \psi_s^*(s_1) \dots \psi_s^*(s_{J-1}) D_\alpha \psi_s(s_J) \psi_s^*(s'_1) \dots \psi_s^*(s'_{J'-1}) (D')^\alpha \psi_s^*(s'_{J'}) \quad (179)$$

where  $\tilde{s}_1 := s$  and the  $*$  subscript denotes an integration of the variables  $\tilde{s}_2, \dots, \tilde{s}_{\tilde{J}}, s_1, \dots, s_J, s'_1, \dots, s'_{J'}$  over the region

$$s = \tilde{s}_1 \leq \dots \leq \tilde{s}_{\tilde{J}} \leq s_1 \leq \dots \leq s_J; \quad \tilde{s}_{\tilde{J}} \leq s'_1 \leq \dots \leq s'_{J'}$$

and we adopt the convention

$$\psi_s^*(\tilde{s}_{\tilde{J}}) \psi_s^*(s_1) \dots \psi_s^*(s_{J-1}) D_\alpha \psi_s^*(s_J) := D_\alpha \psi_s^*(\tilde{s}_{\tilde{J}})$$

if  $J' = 0$ . Thus for instance

$$\begin{aligned} G_{1,0,2}^{\partial,\partial} &= \int_{s'_2 \geq s'_1 \geq s} \partial_\alpha \psi_s^*(s) \psi_s^*(s'_1) \partial^\alpha \psi_s^*(s'_2) \, ds'_1 ds'_2 \\ G_{1,1,1}^{A^*,\partial} &= \int_{s_1, s'_1 \geq s} \psi_s^*(s) A_\alpha^*(s_1) \psi_s^*(s_1) \partial^\alpha \psi_s^*(s'_1) \, ds_1 ds'_1 \\ G_{2,0,2}^{A^*,A^*} &= \int_{s'_2 \geq s'_1 \geq \tilde{s}_2 \geq s} \psi_s^*(s) A_\alpha^*(\tilde{s}_2) \psi_s^*(\tilde{s}_2) \psi_s^*(s'_1) (A^*)^\alpha(s'_2) \psi_s^*(s'_2) \, d\tilde{s}_2 ds'_1 ds'_2 \end{aligned}$$

and so forth. As it will turn out, the  $G_{\tilde{J}, J, J'}^{\partial,\partial}$  expressions will be handled using the trilinear null form estimates (127), (128), whereas the other expressions can be handled using Strichartz estimates as soon as  $\tilde{J} + J + J' \geq 5^{12}$ .

Repeated use of (44), (43) gives the decompositions

$$\begin{aligned} G_1 &= \mathcal{O}(G_{1,0,2}^{\partial,\partial}) + \mathcal{O}(G_{1,0,4}^{\partial,\partial}) + \mathcal{O}(G_{1,0,4}^{\partial,A^*}) \\ G_3 &:= \mathcal{O}(G_{1,2,2}^{\partial,\partial}) + \mathcal{O}(G_{1,2,2}^{\partial,A^*}) + \mathcal{O}(G_{1,2,2}^{A^*,\partial}) + \mathcal{O}(G_{1,2,2}^{A^*,A^*}) \\ G_4 &:= \mathcal{O}(G_{1,1,1}^{\partial,\partial}) + \sum_{(i,j)=(1,3),(3,1),(3,3)} \mathcal{O}(G_{1,i,j}^{\partial,\partial}) + \mathcal{O}(G_{1,i,j}^{\partial,A^*}) + \mathcal{O}(G_{1,i,j}^{A^*,\partial}) + \mathcal{O}(G_{1,i,j}^{A^*,A^*}). \end{aligned}$$

The  $G_2$  term is handled slightly differently, as both of the null form derivatives are falling on the same term (cf. [20, Step 2(d)]). From (43) and the Leibniz rule (4) we have

$$G_2 = \int_s^\infty \mathcal{O}(\psi_s^*(s) (\partial^\alpha \psi_s^*(s')) \psi_s^*(s')) + \mathcal{O}(\psi_s^*(s) \psi_s^*(s') \partial^\alpha \psi_s^*(s')) \, ds'.$$

The second term can be expanded using (44), (43) to take the form

$$\mathcal{O}(G_{2,0,1}^{\partial,\partial}) + \mathcal{O}(G_{2,0,3}^{\partial,\partial}) + \mathcal{O}(G_{2,0,3}^{\partial,A^*}).$$

---

<sup>12</sup>The terms  $G_{\tilde{J}, J, J'}^{\partial,\partial}$  with  $\tilde{J} + J + J' \geq 7$  or  $G_{\tilde{J}, J, J'}^{A^*,A^*}$  with  $\tilde{J} + J + J' \geq 3$  could also be handled by Strichartz estimates, but we will not need to use this fact.

For the first term, we can use the wave-tension field (40) to write  $\partial^\alpha \psi_\alpha^* = w^* + \mathcal{O}((A^*)^\alpha \psi_\alpha^*)$ . Expanding the latter out using (44), (43), we thus see that we can write this contribution to  $G_2$  as

$$\mathcal{O}(H) + \mathcal{O}(G_{2,1,2}^{\partial,\partial}) + \mathcal{O}(G_{2,1,2}^{A^*,\partial}) + \mathcal{O}(G_{2,1,2}^{\partial,A^*}) + \mathcal{O}(G_{2,1,2}^{A^*,A^*})$$

where

$$H := \int_s^\infty \psi_s^*(s) w^*(s') \psi_s^*(s') ds'. \quad (180)$$

Putting this all together, we see that to prove (178), it suffices to show the *null form estimate*

$$\|\partial_x^j P_k G_{\tilde{J},J,J'}^{\partial,\partial}\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{\tilde{J},J,J,\varepsilon} \mu^{2-\varepsilon} c(s) s^{-(j+2)/2} \chi_{k \geq k(s)}^{\delta_2/10} \quad (181)$$

for  $\tilde{J}, J' \geq 1$  and  $J \geq 0$  with  $\tilde{J} + J + J' \geq 3$ , the *Strichartz estimate*

$$\|\partial_x^j P_k G_{\tilde{J},J,J'}^{D,D'}\|_{N_k(I \times \mathbf{R}^2)} \lesssim_{\tilde{J},J,J,\varepsilon} \mu^{2-\varepsilon} c(s) s^{-(j+2)/2} \chi_{k \geq k(s)}^{\delta_2/10} \quad (182)$$

for  $(D, D') = (\partial, A^*), (A^*, \partial), (A^*, A^*)$ ,  $\tilde{J}, J' \geq 1$  and  $J \geq 0$  with  $\tilde{J} + J + J' \geq 5$ , and the *exceptional estimate*

$$\|\partial_x^j P_k H\|_{N_k(I \times \mathbf{R}^2)} \lesssim_\varepsilon \mu^{2-\varepsilon} c(s) s^{-(j+2)/2} \chi_{k \geq k(s)}^{\delta_2/10}. \quad (183)$$

We begin with (183). By (180), the Leibniz rule (4), and Minkowski's inequality, we can bound the left-hand side of (183) by

$$\lesssim \sum_{j_1, j_2, j_3 \geq 0: j_1 + j_2 + j_3 = j} \int_s^\infty \|P_k(\partial_x^{j_1} \psi_s^*(s) \partial_x^{j_2} w^*(s') \partial_x^{j_3} \psi_s^*(s'))\|_{N_k(I \times \mathbf{R}^2)} ds'.$$

Applying (126) and (119), we can bound this by

$$\lesssim \sum_{j_1, j_2, j_3 \geq 0: j_1 + j_2 + j_3 = j} \int_s^\infty \sum_{k_2} \chi_{k \geq k_2}^{\delta_2} \chi_{k = \max(k(s), k_2)}^{\delta_2} \|\partial_x^{j_1} \psi_s^*(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \|\partial_x^{j_2} P_{k_2} w^*(s')\|_{N_{k_2}(I \times \mathbf{R}^2)} \|\partial_x^{j_3} \psi_s^*(s')\|_{N_{k_2}(I \times \mathbf{R}^2)} ds'.$$

Applying (169), (71), we can bound this

$$\lesssim \mu^{2-\varepsilon} s^{-(j+2)/2} \int_s^\infty \sum_{k_2} \chi_{k \geq k_2}^{\delta_2} \chi_{k = \max(k(s), k_2)}^{\delta_2} c(s) c(s')^2 \chi_{k_2 = k(s')}^{\delta_2/10} \frac{ds'}{s'}.$$

Performing the  $k_2$  summation and then using (143), we obtain the claim (183).

We remark that the hardest case here was  $j = 0$ ; the case of higher  $j$  are at least as easy as those of  $j = 0$ , as each derivative will eventually pull out a factor of  $(s')^{-1/2}$  for some  $s' \geq s$ , which is acceptable. Hence, to simplify the notation somewhat, when discussing the other two estimates (181), (182) we shall only discuss the  $j = 0$  case.

We begin with the  $j = 0$  case of (182). We fix  $\tilde{J}, J, J'$  and allow implied constants to depend on these quantities. By (120) and Bernstein's inequality, it suffices to show that

$$\|G_{\tilde{J}, J, J'}^{D, D'}\|_{L_t^1 L_x^1(I \times \mathbf{R}^2)} \lesssim \mu^{2-\varepsilon} c(s)$$

for  $(D, D') = (\partial, A^*), (A^*, \partial), (A^*, A^*)$ .

Suppose first that  $(D, D') = (\partial, A^*)$  or  $(D, D') = (A^*, \partial)$ . Then we see that the integrand in (179) consists of  $J + J' + J'' + 1 \geq 6$  factors, most of which are of  $\psi_s$  type, but with one factor of type  $\partial_{t,x} \psi_s^*$  and one of type  $A_{t,x}^*$ . We use Hölder's inequality and place  $A_{t,x}^*$  in  $L_t^{5/2} L_x^\infty$ , the  $\psi_s^*$  or  $\partial_{t,x} \psi_s^*$  factor with the lowest value of  $s$  in  $L_t^\infty L_x^2$ , the next three lowest in  $L_t^4 L_x^\infty$ , and the remaining factors in  $L_t^\infty L_x^\infty$ . Using Lemma 9.7 and (143), we can bound the  $L_t^1 L_x^1(I \times \mathbf{R}^2)$  norm of the integrand of (179) by

$$\lesssim \mu^{2-\varepsilon} c(s) (s/s_{\max})^\sigma \frac{1}{\tilde{s}_2} \cdots \frac{1}{\tilde{s}_{\tilde{J}}} \frac{1}{s_1} \cdots \frac{1}{s_J} \frac{1}{s'_1} \cdots \frac{1}{s'_{J'}}$$

for some absolute constant  $\sigma > 0$ , where  $s_{\max} := \max(\tilde{s}_1, \dots, \tilde{s}_{\tilde{J}}, s_1, \dots, s_J, s'_1, \dots, s'_{J'})$ ; the point here is that the  $\partial_{t,x}$  and  $A_{t,x}^*$  terms are at high values of  $s$ , leading to the crucial  $(s/s_{\max})^\sigma$  gain. Integrating this using Minkowski's inequality we obtain the claim (182). The case  $(D, D') = (A^*, A^*)$  is similar; the  $\partial_{t,x}$  term has been replaced by an additional  $A_{t,x}^*$  term, but this can be estimated in  $L_t^\infty L_x^\infty$  using Lemma 9.7 to yield the same type of estimates as before.

Finally, we show the  $j = 0$  case of (181). We again fix  $\tilde{J}, J, J'$  and allow implied constants to depend on these quantities.

Let us first handle the case  $(\tilde{J}, J) \neq (1, 0)$ . Here, the integrand in (179) takes the form

$$\Phi \partial^\alpha \psi_s^*(s_J) \partial_\alpha \psi_s^*(s'_{J'}) \quad (184)$$

where  $\Phi$  is the product of many factors of  $\psi_s^*$  evaluated at various  $s' \geq s$ , including one factor of  $\psi_s^*$  itself. By (71), (10), (143), and many applications of (119), we see that

$$\|\Phi\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim c(s) s^{-1} \frac{1}{\tilde{s}_2} \cdots \frac{1}{\tilde{s}_{\tilde{J}}} \frac{1}{s_1} \cdots \frac{1}{s_{J-1}} \frac{1}{s'_1} \cdots \frac{1}{s'_{J'-1}}.$$

Applying (128) and Corollary 9.6, (143) we thus conclude that the  $N_k(I \times \mathbf{R}^2)$  norm of  $P_k$  applied to (184) is bounded by

$$\lesssim \chi_{k=k(s)}^{\delta_2} (s/s_{\max})^\sigma \mu^{2-\varepsilon} c(s) s^{-1} \frac{1}{\tilde{s}_2} \cdots \frac{1}{\tilde{s}_{\tilde{J}}} \frac{1}{s_1} \cdots \frac{1}{s_J} \frac{1}{s'_1} \cdots \frac{1}{s'_{J'}}$$

for some  $\sigma > 0$  (depending on  $\delta_1$ ); integrating this we obtain the claim (181).

The case when  $(\tilde{J}, J) = (1, 0)$  (so  $J' \geq 2$ ) is slightly more delicate, due to the derivative falling on the roughest term. Here, the integrand in (179) takes the form

$$\partial^\alpha \psi_s^*(s) \Phi \partial_\alpha \psi_s^*(s'_{J'}) \quad (185)$$

where  $\Phi := \psi_s^*(s'_1) \dots \psi_s^*(s'_{J'-1})$ . By arguing as before, we have

$$\|\Phi\|_{S_{k(s'_1)}(I \times \mathbf{R}^2)} \lesssim c(s'_1) \frac{1}{s'_1} \dots \frac{1}{s'_{J'-1}}.$$

Applying (128) and Corollary 9.6 (but now refraining from using (143) to eliminate some key  $c()$  factors), the  $N_k(I \times \mathbf{R}^2)$  norm of  $P_k$  applied to (185) is bounded by

$$\lesssim \chi_{k=k(s)}^{\delta_1/2} (s'_1/s'_{J'})^{\delta_1/4} \mu^{2-\varepsilon} c(s) c(s'_1) c(s'_{J'}) s^{-1} \frac{1}{s'_1} \dots \frac{1}{s'_{J'}}.$$

We can use (10) to bound  $(s'_1/s'_{J'})^{\delta_1/4} c(s'_1) c(s'_{J'})$  by  $c(s'_1)^2 (s'_1/s'_{J'})^{\delta_1/4}$ . Integrating out  $s'_{J'}$ , then  $s'_{J'-1}$ , and so forth down to  $s'_1$  (using (143) to control the final integral) we obtain the claim (181).

The proof of (177) follows by applying (5) to all the expressions above and repeating the argument in the obvious manner; we omit the details.  $\square$

Now, at long last, we are ready to prove (144). Fix  $0 \leq j \leq 10$  and  $s > 0$ . Applying Lemma 9.10 (with  $j = 11$ , say), Littlewood-Paley decomposition, and the triangle inequality, we have

$$s^{1+\frac{j}{2}} \|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim_\varepsilon \sum_{k'} \chi_{k' \leq k(s)}^{-10} s \|P_{k'} \partial_{t,x} \psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} + \mu^{1-\varepsilon} \delta c(s).$$

Applying Lemma 7.6 we conclude

$$s^{1+\frac{j}{2}} \|\partial_x^j \delta \psi_s(s)\|_{S_{k(s)}(I \times \mathbf{R}^2)} \lesssim \delta c_0(s) + \mu^{1-\varepsilon} \delta c(s)$$

for some frequency envelope  $\delta c_0$  of energy  $O(d_{\mathcal{H}^1}(\Psi_{s,t,x}, \Psi'_{s,t,x})^2)$ .

Applying (142), (134) we conclude

$$\delta c(s) \lesssim \delta c_0(s) + \mu^{1-\varepsilon} \delta c(s).$$

If  $\mu$  is sufficiently small, we thus conclude

$$\delta c(s) \lesssim \delta c_0(s)$$

and (144) follows from the energy bound on  $\delta c_0$ . The proof of (58) is now (finally!) complete.

**9.12. Return to the energy metric.** We now establish part (iv) of Theorem 5.11, in which we control the energy metric by the  $S_\mu^1$  metric.

Fix  $I, \mu, M, E$ ; we let all implied constants depend on  $M, E$ . Suppose  $\Psi_{s,t,x}, \Psi'_{s,t,x} \in \text{WMC}(I, E)$  is such that  $\|\Psi_{s,t,x}\|_{S_\mu^1(I)}, \|\Psi'_{s,t,x}\|_{S_\mu^1(I)} \leq M$ . Let us write

$$\sigma := d_{S_\mu^1, I}(\Psi_{s,t,x}, \Psi'_{s,t,x})$$

and let  $t \in I$  be fixed; we omit the explicit mention of this parameter. By (56), our task is to show that

$$\int_0^\infty \|\delta\psi_s(s)\|_{L_x^2(\mathbf{R}^2)}^2 ds \lesssim \sigma^2$$

and

$$\|\delta\psi_t(0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \sigma.$$

Let  $c, \delta c$  be as in the preceding section, then  $\delta c$  has energy  $O(\sigma^2)$  by construction, while  $c$  has energy  $O(1)$ . From Lemma 9.7 (which does not require a smallness hypothesis on  $\mu$ ), we have the bounds

$$\begin{aligned} \|\partial_x^j \psi_s^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c(s) s^{-(j+1)/2} \\ \|\partial_x^j \psi_s^*(s)\|_{L_x^\infty(\mathbf{R}^2)} &\lesssim_j c(s) s^{-(j+2)/2} \\ \|\partial_x^{j+1} \psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} &\lesssim_j c(s) s^{-(j+1)/2} \\ \|\partial_x^j \psi_{t,x}^*(s)\|_{L_x^\infty(\mathbf{R}^2)} &\lesssim_j c(s) s^{-(j+1)/2} \end{aligned}$$

for all  $j \geq 0$ , and similarly for  $\delta\psi_s$  and  $\delta\psi_{t,x}$  with  $c$  replaced by  $\delta c$ . Also, from [25, Lemma 7.2] we have

$$\|\psi_{t,x}^*(s)\|_{L_x^2(\mathbf{R}^2)} \lesssim 1.$$

Repeating the arguments used to prove Lemma 7.7, we see that

$$\|P_k \delta\psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim_j \delta c_0(s) s^{-1/2} \chi_{k \leq k(s)}^j \chi_{k=k(s)}^{0.1} \quad (186)$$

and

$$\|P_k \delta\psi_t(0, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \delta c_0(2^{-2k}), \quad (187)$$

for all  $s > 0$ ,  $j \geq 0$ , and  $k \in \mathbf{Z}$ , and the claim follows.

**9.13. Persistence of regularity.** We now prove (57). We allow all implied constants to depend on  $M$  and  $\|\phi[t_0]\|_{\mathcal{H}^{100}}$ , in particular we have  $\|\phi\|_{S_\mu^1(I \times \mathbf{R}^2)} \lesssim 1$  and hence  $E(\phi) \lesssim 1$ . Our task is now to show that if  $\mu$  is sufficiently small, then

$$\|\phi[t]\|_{\mathcal{H}_{\text{loc}}^{1+\delta_0/2}} \lesssim_\mu 1. \quad (188)$$

for all  $t \in I$ .



The first step is to show that some of the  $\mathcal{H}^{100}$  regularity persists under the heat flow. Let  $\Psi_{s,t,x} \in \text{WMC}(\phi, I)$  be one of the differentiated fields for  $\phi$  (the exact choice does not matter).

**Lemma 9.14** (Smooth control of  $\psi_s$ ). *For any  $0 < s \lesssim 1$  we have*

$$\|\psi_s(s, t_0)\|_{H_x^{90}(\mathbf{R}^2)} + \|\partial_{t,x}\psi_s(s, t_0)\|_{H_x^{90}(\mathbf{R}^2)} \lesssim 1.$$

*Proof.* By hypothesis,

$$\|\phi[t_0]\|_{\mathcal{H}^{100}} \lesssim 1.$$

From (18), (19), for each  $x_0$  one can find  $U_{x_0} \in SO(m, 1)$  such that

$$\|\eta(\cdot - x_0)U_{x_0}(\phi_0)\|_{H_x^{100}(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U_{x_0}(\phi_1)\|_{H_x^{99}(\mathbf{R}^2)} \lesssim 1$$

which by Sobolev embedding implies that  $U_{x_0}(\phi_0) = O(1)$  on  $B(x_0, 1)$ . From this and many applications of the Leibniz rule (4), we conclude that

$$\begin{aligned} & \| |(\phi^* \nabla_x)^j \partial_x \phi_0|_{\phi^* h} \|_{L^2(B(x_0, 1))} + \| |(\phi^* \nabla_x)^j \phi_1|_{\phi^* h} \|_{L^2(B(x_0, 1))} \\ & \lesssim \inf_{U \in SO(m, 1)} \|\eta(\cdot - x_0)U(\phi_0)\|_{H_x^s(\mathbf{R}^2)} + \|\eta(\cdot - x_0)U(\phi_1)\|_{H_x^{s-1}(\mathbf{R}^2)} \end{aligned}$$

for all  $x_0 \in \mathbf{R}^2$  and  $0 \leq j \leq 99$ . Integrating this over  $x_0$  and using (19), we conclude that

$$\| |(\phi^* \nabla_x)^j \partial_x \phi_0|_{\phi^* h} \|_{L^2(\mathbf{R}^2)} + \| |(\phi^* \nabla_x)^j \phi_1|_{\phi^* h} \|_{L^2(\mathbf{R}^2)} \lesssim 1$$

and thus in the caloric gauge

$$\|D_x^j \psi_{t,x}(0, t_0)\|_{L^2(\mathbf{R}^2)} \lesssim 1$$

for  $0 \leq j \leq 99$ .

For  $s \geq 0$ , define

$$E_{100}(s) := \sum_{j=0}^{99} \int_{\mathbf{R}^2} |D_x^j \psi_{t,x}(s, t_0)|^2 dx$$

thus  $E_{100}(0) = O(1)$ . On the other hand, from Theorem 7.2 we can find an envelope  $c_0$  of energy  $O(1)$  such that  $D_x^j \psi_{t,x}(s, t_0) = O(c_0(s)s^{-(j+1)/2})$  for all  $0 \leq j \leq 100$ . Using this, (46), and many applications of covariant integration by parts, one obtains the differential inequality

$$\partial_s E_{100}(s) = O(c_0(s)^2 s^{-1} E_{100}(s))$$

which by Gronwall's inequality and the fact that  $c_0$  has energy  $O(1)$  implies that  $E_{100}(s) = O(1)$  for all  $s$ , thus

$$\|D_x^j \psi_{t,x}(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim 1$$

for all  $0 \leq j \leq 99$  and  $s > 0$ . Using the covariant Gagliardo-Nirenberg inequality

$$\|\varphi\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \|\varphi\|_{L_x^2(\mathbf{R}^2)}^{1/3} \|D_x \varphi\|_{L_x^2(\mathbf{R}^2)}^{1/3} \|D_x^2 \varphi\|_{L_x^2(\mathbf{R}^2)}^{1/3}$$

(see the proof of [25, Lemma 4.2]) we conclude that

$$\|D_x^j \psi_{t,x}(s, t_0)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim 1$$

for all  $0 \leq j \leq 97$  and  $s > 0$ . Combining this with what one gets from Theorem 7.2, we conclude

$$\|D_x^j \psi_{t,x}(s, t_0)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \langle s \rangle^{-(j+1)/2}$$

for all  $0 \leq j \leq 97$  and  $s > 0$ . In particular, by (38) this implies

$$\|D_x^j \psi_s(s, t_0)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \langle s \rangle^{-(j+2)/2}$$

for all  $0 \leq j \leq 96$  and  $s > 0$ . Similar reasoning gives

$$\|D_x^j \psi_{t,x}(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \langle s \rangle^{-j/2}$$

and

$$\|D_x^j \psi_s(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \langle s \rangle^{-(j+1)/2}$$

for this range. Applying (43), the covariant Leibniz rule, and Minkowski's inequality, we conclude

$$\|D_x^j A_{t,x}(s, t_0)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim \langle s \rangle^{-(j+1)/2}$$

and

$$\|D_x^j A_{t,x}(s, t_0)\|_{L_x^2(\mathbf{R}^2)} \lesssim \langle s \rangle^{-j/2}$$

for all  $0 \leq j \leq 96$  and  $s > 0$ , where  $D_x B := \partial_x B + [A_x, B]$  for any matrix field  $B$ . (The  $j = 0$   $L^2$  bound on  $A_{t,x}$  was already established in [25, Proposition 4.3].) Writing non-covariant derivatives in terms of covariant ones, we conclude that the above two estimates continue to hold if  $D_x$  is replaced by  $\partial_x$ . We can then similarly replace  $D_x$  by  $\partial_x$  in all previous estimates. We conclude that

$$\|\Psi_{t,x}(s, t_0)\|_{H_x^{96}(\mathbf{R}^2)} + \|\psi_s(s, t_0)\|_{H_x^{96}(\mathbf{R}^2)} \lesssim 1$$

for all  $0 \leq s \lesssim 1$ . The claim then easily follows from (38) and (36).  $\square$

Let  $c(s)$  be the envelope from the proof of (58). We can obtain a decay estimate:

**Lemma 9.15** (Breaking the scaling barrier). *We have*

$$c(s) \lesssim_\mu \min(1, s^{-\delta_0/2})$$

for all  $s > 0$ .

*Proof.* Let  $c(s)$  be the envelope from the proof of (26). From Lemma 9.14 and (176) we have

$$\|P_k \psi_s^*(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim_\varepsilon \min(2^k, 2^{-90k}) + \mu^{2-\varepsilon} c(s) s^{-1} \chi_{k \geq k(s)}^{\delta_2/10} \chi_{k \leq k(s)}^{11}$$

for  $0 < s \lesssim 1$  and  $\varepsilon > 0$ . This implies that

$$s^{1+\frac{j}{2}} \|\partial_x^j \psi_s(s)\|_{S_k(I \times \mathbf{R}^2)} \lesssim_\varepsilon s + \mu^{2-\varepsilon} c(s)$$

for all  $0 \leq j \leq 10$ ,  $0 \leq s \lesssim 1$ , and  $\varepsilon$ ; by (117), (140) and (10) this implies that

$$d_k(0, \Psi) \lesssim_\varepsilon \mu^{-1} 2^{-k} + \mu^{1-\varepsilon} c(2^{-2k})$$

for  $k \geq 0$ . For  $k < 0$  we have the trivial bound  $d_k(0, \Psi) \lesssim c(2^{-2k})$  (by (141)); we can unify these bounds to obtain

$$d_k(0, \Psi) \lesssim_\varepsilon \mu^{-1} \langle 2^k \rangle^{-1} + (\mu^{1-\varepsilon} + \langle 2^k \rangle^{-1}) c(2^{-2k})$$

for all  $k$ . Inserting this into (141) and using (10) we obtain

$$c(2^{-2k}) \lesssim_\varepsilon \mu^{-1} \langle 2^k \rangle^{-\delta_0} + (\mu^{1-\varepsilon} + \langle 2^k \rangle^{-\delta_0}) c(2^{-2k})$$

and thus (if  $\mu$  is small enough, and  $\varepsilon$  is chosen to be, say,  $1/2$ ) we see that

$$c(2^{-2k}) \lesssim_\mu \langle 2^k \rangle^{-\delta_0}$$

for all  $k$  (note that this already follows from (143) if  $2^{-k} \lesssim 1$ ). The claim follows.  $\square$

We now fix a time  $t \in I$ . From Lemma 9.15 and Lemma 9.7 one has

$$\|\Psi_{t,x}(s, t)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim_\mu \min(1, s^{-\delta_0/2}) s^{-1/2}$$

and

$$\|\psi_s(s, t)\|_{L_x^\infty(\mathbf{R}^2)} \lesssim_\mu \min(1, s^{-\delta_0/2}) s^{-1}$$

for all  $s > 0$ . Converting this back to the original wave map  $\phi$ , we see that

$$|\partial_{t,x} \phi(s, t, x)|_{\phi^* h} \lesssim_\mu \min(1, s^{-\delta_0/2}) s^{-1/2}$$

and

$$|\partial_s \phi(s, t, x)|_{\phi^* h} \lesssim_\mu \min(1, s^{-\delta_0/2}) s^{-1}$$

for all  $s > 0$  and  $x \in \mathbf{R}^2$ . Observe that the right-hand sides here are locally integrable in  $s$ . By the fundamental theorem of calculus on  $\mathbf{H}$ , we conclude that

$$d_{\mathbf{H}}(\phi(s_1, t, x_1), \phi(s_2, t, x_2)) \lesssim_\mu 1$$

whenever  $0 < s_1, s_2 \lesssim 1$  and  $x_1, x_2 \in \mathbf{R}^2$  are such that  $|x_2 - x_1| \lesssim 1$ , and  $d_{\mathbf{H}}$  is the distance function on  $\mathbf{H}$  induced by the metric  $h$ . (In fact we even obtain some Hölder continuity estimates here, but we will not exploit this.)

From Lemma 9.15 and Lemma 9.7 we have

$$\|\partial_x^j \psi_s\|_{L_x^2(\mathbf{R}^2)} \lesssim_{\mu,j} s^{-\delta_0/2} s^{-(j+1)/2}$$

for all  $0 < s \leq 1$  and  $j > 0$ . By interpolation this implies that

$$\|\psi_s(s)\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} \lesssim_\mu s^{-\delta_0/4} s^{-1} \quad (189)$$

for all  $0 < s \leq 1$ . Similarly we have

$$\|\partial_x^j \partial_t \psi_s\|_{L_x^2(\mathbf{R}^2)} \lesssim_{\mu,j} s^{-\delta_0/2} s^{-(j+2)/2}$$

and hence

$$\|\partial_t \psi_s(s)\|_{H_x^{\delta_0/2}(\mathbf{R}^2)} \lesssim_\mu s^{-\delta_0/4} s^{-1}. \quad (190)$$

Let  $\frac{1}{100}\mathbf{Z}^2 \subset \mathbf{R}^2$  be the standard lattice with spacing  $\frac{1}{100}$  in the spatial domain  $\mathbf{R}^2$ . For each  $x_0 \in \frac{1}{100}\mathbf{Z}^2$ , let  $U_{x_0} \in SO(m, 1)$  be a Lorentz rotation that sends  $\phi(1, t, x_0)$  to the origin  $(1, 0) \in \mathbf{H}$ . We write  $\phi_{x_0} := U_{x_0}(\phi)$ ,  $e_{x_0} := U_{x_0} \circ e$ . By the preceding discussion, we see that

$$\phi_{x_0}(s, t, x_1) = O(1)$$

whenever  $x_1 = x_0 + O(1)$  and  $0 \leq t \leq 1$ ; if we set  $\Phi_{x_0} := (\phi_{x_0}, e_{x_0})$ , this implies that

$$\Phi_{x_0}(s, t, x_1) = O(1) \quad (191)$$

for the same range of  $t, x_1$ . Also, from Lemma 9.7 we have

$$|\partial_x^j \Psi_{t,x}(1, t, x)| \lesssim_{\mu,j} 1 \quad (192)$$

for  $j \geq 0$  and  $x \in \mathbf{R}^2$ . From (34), (33) (applied to the rotated wave map  $\phi_{x_0}$  and its associated rotated frame  $e_{x_0}$ ) we have the differential equations

$$\partial_{t,x} \phi_{x_0} = e_{x_0}(\psi_{t,x})$$

and

$$(\phi_{x_0}^* \partial_{t,x})(e_{x_0})_a = (A_{t,x})_{ab} e_b.$$

We write these equations schematically using (14) as

$$\partial_{t,x} \Phi_{x_0} = \mathcal{O}(\Phi_{x_0} \Psi_{t,x}) + \mathcal{O}(\Phi_{x_0}^2 \Psi_{t,x}) \quad (193)$$

where  $\Phi_{x_0} := (\phi_{x_0}, e_{x_0})$ . From this, (191), (192), and many applications of the Leibniz rule (4) and Gronwall's inequality, we thus conclude that

$$|\partial_x^j \Phi_{x_0}(1, t, x)|, |\partial_x^j \partial_t \Phi_{x_0}(1, t, x)| \lesssim_j 1 \quad (194)$$

whenever  $j \geq 0$  and  $|x - x_0| \lesssim 1$ .

For each  $0 \leq s \leq 1$ , define

$$f(s) := \sup_{x_0 \in \frac{1}{100}\mathbf{Z}^2} \|\eta(\cdot - x_0) \Phi_{x_0}(s, t)\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} + \|\eta(\cdot - x_0) \partial_t \Phi_{x_0}(s, t)\|_{H_x^{\delta_0/2}(\mathbf{R}^2)}.$$

From (194) we see that

$$f(1) \lesssim 1. \quad (195)$$

We shall shortly show that

$$\limsup_{ds \rightarrow 0} \frac{|f(s+ds) - f(s)|}{|ds|} \lesssim_\mu s^{-\delta_0/4} s^{-1} f(s) \quad (196)$$

for all  $0 < s \leq 1$ . Note that  $s^{-\delta_0/4} s^{-1}$  is integrable on this interval. From this, Gronwall's inequality and (195) we see that  $f(s) \lesssim 1$  for all  $0 < s \leq 1$ ; taking limits as  $s \rightarrow 0$  we conclude  $f(0) \lesssim 1$ , from which (188) (and hence (26)) easily follows. Thus the only remaining task is to show (196).

By Minkowski's inequality and the fundamental theorem of calculus (and the smoothness of  $\phi$ ) it suffices to show that

$$\|\eta(\cdot - x_0)\partial_s\Phi_{x_0}(s, t)\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} + \|\eta(\cdot - x_0)\partial_t\partial_s\Phi_{x_0}(s, t)\|_{H_x^{\delta_0/2}(\mathbf{R}^2)} \lesssim_\mu s^{-\delta_0/4} s^{-1} f(s)$$

for all  $0 < s \leq 1$  and  $x_0 \in \frac{1}{100}\mathbf{Z}^2$ .

Fix  $x_0, s$ . By repeating the derivation of (193) we have

$$\partial_s\Phi_{x_0} = \mathcal{O}(\Phi_{x_0}\psi_s) + \mathcal{O}(\Phi_{x_0}^2\psi_s); \quad (197)$$

using the standard Sobolev product estimates

$$\|uv\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} \lesssim \|u\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} \|v\|_{L_x^\infty(\mathbf{R}^2)} + \|u\|_{L_x^\infty(\mathbf{R}^2)} \|v\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)}$$

(see e.g. [22, Lemma A.8] for a proof) together with (189) and (191), we conclude that

$$\|\eta(\cdot - x_0)\partial_s\Phi_{x_0}(s, t)\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} \lesssim_\mu s^{-\delta_0/4} s^{-1} f(s).$$

Similarly, applying  $\partial_t$  to (197) and using the above estimates and the product estimate

$$\|uv\|_{H_x^{\delta_0/2}(\mathbf{R}^2)} \lesssim \|u\|_{H_x^{1+\delta_0/2}(\mathbf{R}^2)} \|v\|_{H_x^{\delta_0/2}(\mathbf{R}^2)}$$

we obtain

$$\|\eta(\cdot - x_0)\partial_t\partial_s\Phi_{x_0}(s, t)\|_{H_x^{\delta_0/2}(\mathbf{R}^2)} \lesssim_\mu s^{-\delta_0/4} s^{-1} f(s).$$

The claim (196) (and hence (57)) follows.

## 10. CONSTRUCTION OF FUNCTION SPACES

We now prove Theorem 8.1. The spaces  $S_k, S_{\mu,k}, N_k$  we need to employ are quite complicated. Fortunately, we can take their definition directly from [20], with few modifications, so that the verification of Theorem 8.1 largely consists of citing the relevant results from that paper. (The spaces in [20] were in turn based on those in [28].)

More precisely, we will take  $S_k, N_k$  to be the spaces introduced in [20, Theorem 3], and constructed in [20, Section 10]. The exact definition of these spaces is complicated, but will not actually be needed for this paper, as we can largely deduce everything we need from [20, Theorem 3]. For instance:

- The product estimate (119) follows immediately from [20, Equation (125)].
- The estimate (120) is precisely [20, Equation (25)].
- The estimate (121) follows immediately from [20, Equation (87)].
- The estimate (123) is precisely [20, Equation (26)].
- The estimate (124) is precisely [20, Equation (27)].

- The estimate (126) is precisely [20, Equation (29)] (and some Littlewood-Paley decomposition).
- The estimate (129) is precisely [20, Equation (35)].
- The estimate (131) follows from [10, Lemma 3.1] (or [10, Lemma 6.7]).

We will however point out here one feature of the  $S_k$  spaces that was not emphasised in those papers, and which is needed in order to establish the continuity claim in Theorem 8.1. As noted in passing in [20], the spaces  $S_k$  constructed in that paper enjoy the discrete scale invariance

$$\|\phi^{(j)}\|_{S_{k+j}(2^{-j}I)} = \|\phi\|_{S_k(I \times \mathbf{R}^2)} \quad (198)$$

for all intervals  $I$ ,  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ , and  $k, j \in \mathbf{Z}$ , where  $\phi^{(j)}(t, x) := \phi(2^j t, 2^j x)$ . It is also possible to establish a continuous analogue

$$\|\phi^{(j)}\|_{S_{k+[j]}(2^{-j}I)} \lesssim \|\phi\|_{S_k(I \times \mathbf{R}^2)} \quad (199)$$

of this scale invariance, where  $j$  is now taken to be real instead of integer, and  $[j]$  is the nearest integer to  $j$  (rounding down, say). Indeed, to establish this, one can use (198) to reduce to the case  $-1/2 < j \leq 1/2$ , in which case the claim follows by modifying the proof of [20, Lemma 9]. We omit the details<sup>13</sup>.

Let us now define the modified  $S_k$  norm

$$\|\phi\|_{\overline{S_k}(I \times \mathbf{R}^2)} := \sup_{j \in \mathbf{R}} 2^{-100|j|} \|\phi^{(j)}\|_{S_k(2^{-j}I \times \mathbf{R}^2)}$$

From (199), (121) we see that the  $\overline{S_k}(I \times \mathbf{R}^2)$  norm is equivalent to the  $S_k(I \times \mathbf{R}^2)$  norm up to absolute constants. Furthermore, observe that  $\|\phi^{(j)}\|_{\overline{S_k}(2^{-j}I \times \mathbf{R}^2)}$  varies continuously in  $j \in \mathbf{R}$ . The construction of  $S_k(I \times \mathbf{R}^2)$  in [20] ensures that this norm is monotone in  $I$ , and so the  $\overline{S_k}(I \times \mathbf{R}^2)$  norm is monotone also.

We now define  $S_{\mu,k}(I \times \mathbf{R}^2)$  to be the atomic Banach space whose atoms  $\phi$  are one of two types:

- (Small atoms)  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$  with  $\|\phi\|_{\overline{S_k}(I \times \mathbf{R}^2)} \leq \mu$ .
- (Integrable atoms)  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$  with  $\|\phi\|_{\overline{S_k}(I \times \mathbf{R}^2)} \leq 1$ , and  $\|\partial_{t,x}\phi\|_{L_t^1 L_x^\infty(I \times \mathbf{R}^2)} \leq \mu^5$ .

---

<sup>13</sup>There are several alternate ways to deal with this. One is to adjust the definition of the spaces in [20] so that frequency parameters such as  $k, j$  vary over the reals rather than the integers. Another is to abandon any proof of continuity, and settle instead for the weaker property of quasicontinuity as in [20]; it is not hard to see that quasicontinuity serves as a reasonable substitute for continuity for the purposes of performing continuity arguments.

The estimate (117) is now immediate, as is the monotonicity of  $S_{\mu,k}(I \times \mathbf{R}^2)$  in  $I$ , while (122) follows from (121). Also, these spaces are clearly invariant under space and time translations. Now we establish continuity of  $\|\phi\|_{S_{\mu,k}(I \times \mathbf{R}^2)}$  in  $I$  for  $\phi \in \text{Schwartz}(I \times \mathbf{R}^2)$ , using a scaling argument of Tataru (cf. [10, Lemma 3.2]). By monotonicity and time translation symmetry it will suffice to show that  $\|\phi\|_{S_{\mu,k}([-T,T] \times \mathbf{R}^2)}$  is continuous in  $T$  for  $\phi \in \text{Schwartz}([-T,T] \times \mathbf{R}^2)$ . We can extend  $\phi$  to lie in  $\text{Schwartz}(I \times \mathbf{R}^2)$  for some neighbourhood  $I$  of  $[-T,T]$ . From the continuity of  $\|\phi^{(j)}\|_{\overline{S_k(2^{-j}I)}}$  and  $\|\partial_{t,x}\phi^{(j)}\|_{L_t^1 L_x^\infty(2^{-j}I)}$  in  $j$  (in fact the latter quantity is invariant in  $j$ ) we already know that  $\|\phi^{(j)}\|_{S_{\mu,k}([-2^{-j}T, 2^{-j}T] \times \mathbf{R}^2)}$  is continuous in  $j$ , so by the triangle inequality it suffices to show that

$$\lim_{j \rightarrow 0} \|\phi^{(j)} - \phi\|_{S_{\mu,k}([-2^{-j}T, 2^{-j}T] \times \mathbf{R}^2)} = 0.$$

But from (132), (117), we can bound

$$\begin{aligned} \|\phi^{(j)} - \phi\|_{S_{\mu,k}([-2^{-j}T, 2^{-j}T] \times \mathbf{R}^2)} &\lesssim_\mu \sum_{k'} \chi_{k=k'}^{-\delta_1} [\|P_{k'} \partial_{t,x}(\phi^{(j)} - \phi)(0)\|_{L_x^2(\mathbf{R}^2)} \\ &\quad + \|P_{k'} \square(\phi^{(j)} - \phi)\|_{L_t^1 L_x^2([-2^{-j}T, 2^{-j}T] \times \mathbf{R}^2)}]. \end{aligned}$$

Since  $\phi$  is Schwartz, it is not difficult to see (by a dominated convergence argument) that the right hand side goes to zero as  $j \rightarrow 0$ , and the claim follows.

The estimate (130) follows from (131) (for small atoms) and from interpolating (131) with the  $L_t^1 L_x^\infty$  estimate (for integrable atoms). Now we establish the vanishing property (118). By time translation we may take  $t_0 = 0$ , and then we may extend  $\phi$  slightly so that it lies in  $\text{Schwartz}([- \varepsilon, \varepsilon] \times \mathbf{R}^2)$  for some  $\varepsilon > 0$ . It suffices to show that

$$\limsup_{T \rightarrow 0} \|\phi\|_{S_{\mu,k}([-T,T] \times \mathbf{R}^2)} \lesssim \sum_{k'} \chi_{k=k'}^{-\delta_1} \|\partial_{t,x} P_{k'} \phi(t_0)\|_{L_x^2(\mathbf{R}^2)}.$$

By definition of  $S_{\mu,k}$  and the triangle inequality, we can estimate

$$\|\phi\|_{S_{\mu,k}([-T,T] \times \mathbf{R}^2)} \lesssim \sum_{k'} \|P_{k'} \phi\|_{S_k([-T,T] \times \mathbf{R}^2)} + \frac{1}{\mu^5} \|\partial_{t,x} P_{k'} \phi\|_{L_t^1 L_x^\infty([-T,T] \times \mathbf{R}^2)}.$$

From (132) we have

$$\|P_{k'} \phi\|_{S_k([-T,T] \times \mathbf{R}^2)} \lesssim \chi_{k=k'}^{-\delta_1} [\|P_{k'} \partial_{t,x} \phi(0)\|_{L_x^2(\mathbf{R}^2)} + \|P_{k'} \square \phi\|_{L_t^1 L_x^2([-T,T] \times \mathbf{R}^2)}].$$

Since  $\phi \in \text{Schwartz}([- \varepsilon, \varepsilon] \times \mathbf{R}^2)$ , an easy application of the dominated and monotone convergence theorems shows that

$$\lim_{T \rightarrow 0} \sum_{k'} \|\partial_{t,x} P_{k'} \phi\|_{L_t^1 L_x^\infty([-T,T] \times \mathbf{R}^2)} = 0$$

and

$$\lim_{T \rightarrow 0} \sum_{k'} \chi_{k=k'}^{-\delta_1} \|P_{k'} \square \phi\|_{L_t^1 L_x^2([-T,T] \times \mathbf{R}^2)} = 0,$$

and the claim follows.

To prove (125), we can use exactly the same argument used to prove (61) in Section 7; note from [20, Equation (87)] that the  $S_k(\mathbf{R})$  norm (which the  $S_k(I \times \mathbf{R}^2)$  norm is the restriction of) has a form extremely similar to that of (96). We omit the details.

Finally, we verify (127), (128). The  $\varepsilon = 1$  cases of these inequalities are exactly [20, Equation (31)], so it suffices to establish the  $\varepsilon = 0$  case.

We begin with (127). By symmetry we may take  $\mu_2 = \mu$ . We may reduce to the case when  $\phi^{(2)}$  is an atom. If it is a small atom, the claim again follows from [20, Equation (31)]. If instead  $\phi^{(2)}$  is an integrable atom, we use (120) and Hölder's inequality to estimate

$$\begin{aligned} \|P_k(\phi^{(1)} \partial_\alpha \phi^{(2)} \partial^\alpha \phi^{(3)})\|_{N_k(I \times \mathbf{R}^2)} &\lesssim \|\phi^{(1)} \partial_\alpha \phi^{(2)} \partial^\alpha \phi^{(3)}\|_{L_t^1 L_x^2(I \times \mathbf{R}^2)} \\ &\lesssim \|\phi^{(1)}\|_{L_t^\infty L_x^\infty(I \times \mathbf{R}^2)} \|\partial_{t,x} \phi^{(2)}\|_{L_t^1 L_x^\infty(I \times \mathbf{R}^2)} \|\partial_{t,x} \phi^{(3)}\|_{L_t^\infty L_x^2(I \times \mathbf{R}^2)}, \end{aligned}$$

which is acceptable (with several powers of  $\mu$  to spare) from (131), (129), and the definition of an integrable atom. The proof of (128) is similar; if both  $\phi^{(2)}$  and  $\phi^{(3)}$  are small atoms, one can use [20, Equation (31)], and when there is an integrable atom, one uses Hölder's inequality as above. The proof of Theorem 8.1 is complete.

## REFERENCES

- [1] P. D'Ancona, V. Georgiev, *On the continuity of the solution operator of the wave maps system*, Comm. Pure Appl. Math. **57** (2004), 357–383.
- [2] I. Bejenaru, A. Ionescu, C. Kenig, D. Tataru, *Global Schrödinger maps*, preprint.
- [3] J. Eells, H. Sampson, *Harmonic mappings of Riemannian manifolds*, Amer. J. Math. **86** (1964), 109–160.
- [4] F. Hélein, *Régularité des applications faiblement harmoniques entre une surface et une variété Riemannienne*, C.R. Acad. Sci. Paris Ser. I Math., **312** (1991), 591–596.
- [5] S. Klainerman, *On the regularity of classical field theories in Minkowski space-time  $\mathbf{R}^{3+1}$* , Prog. in Nonlin. Diff. Eq. and their Applic., **29**, (1997), Birkhäuser, 113–150.
- [6] S. Klainerman, *PDE as a unified subject*, Special Volume GAFA 2000, 279–315.
- [7] S. Klainerman, S. Selberg, *Remark on the optimal regularity for equations of wave maps type*, C.P.D.E., **22** (1997), 901–918.
- [8] S. Klainerman, S. Selberg, *Bilinear estimates and applications to nonlinear wave equations*, Commun. Contemp. Math. **4** (2002), 223–295.
- [9] J. Krieger, *Global regularity of wave maps from  $\mathbf{R}^{3+1}$  to surfaces*, Comm. Math. Phys. **238** (2003), 333–366.
- [10] J. Krieger, *Global regularity of wave maps from  $\mathbf{R}^{2+1}$  to  $\mathbf{H}^2$* , Comm. Math. Phys. **250** (2004), 507–580.
- [11] J. Krieger, *Global regularity and singularity development for wave maps.*, preprint.



- [12] L. Lemaire, *Applications harmoniques de surfaces riemanniennes*, J. Diff. Geom. **13** (1978).
- [13] I. Rodnianski, *The wave map problem. Small data critical regularity*, Seminaire Bourbaki, 58eme annee, 2005–2006, no. 965.
- [14] J. Shatah, *The Cauchy problem for harmonic maps on Minkowski space*, in Proceed. Inter. Congress of Math. 1994, Birkhäuser, 1126–1132.
- [15] J. Shatah, M. Struwe, *Geometric Wave Equations*, Courant Lecture Notes in Mathematics **2** (1998)
- [16] C. D. Sogge, *Lectures on Nonlinear Wave Equations*, Monographs in Analysis II, International Press, 1995.
- [17] M. Struwe, *Wave Maps*, in *Nonlinear Partial Differential Equations in Geometry and Physics*, Prog. in Nonlin. Diff. Eq. and their Applic., **29**, (1997), Birkhäuser, 113–150.
- [18] T. Tao, *Low regularity semilinear wave equations*, to appear, Comm. PDE.
- [19] T. Tao, *Global regularity of wave maps I. Small critical Sobolev norm in high dimension*, IMRN **7** (2001), 299–328.
- [20] T. Tao, *Global regularity of wave maps II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), 443–544.
- [21] T. Tao, *Geometric renormalization of large energy wave maps*, Journees “Equations aux derives partielles”, Forges les Eaux, 7–11 June 2004, XI 1–32.
- [22] T. Tao, *Nonlinear dispersive equations. Local and global analysis*. CBMS Regional Conference Series in Mathematics, **106**. American Mathematical Society, Providence, RI, 2006.
- [23] T. Tao, *Global behaviour of nonlinear dispersive and wave equations*, Current Developments in Mathematics 2006, International Press. 255–340.
- [24] T. Tao, *Global regularity of wave maps III. Large energy from  $\mathbf{R}^{1+2}$  to hyperbolic spaces*, preprint.
- [25] T. Tao, *Global regularity of wave maps IV. Absence of stationary or self-similar solutions in the energy class*, in preparation.
- [26] T. Tao, *Global regularity of wave maps VI. Abstract theory of minimal-energy blowup solutions*, in preparation.
- [27] T. Tao, *Global regularity of wave maps VII. Control of delocalised or dispersed solutions*, in preparation.
- [28] D. Tataru, *On global existence and scattering for the wave maps equation*, Amer. J. Math. **123** (2001), no. 1, 37–77.
- [29] D. Tataru, *Rough solutions for the wave maps equation*, Amer. J. Math. **127** (2005), no. 2, 293–377.
- [30] D. Tataru, *The wave maps equation*, Bull. Amer. Math. Soc. (N.S.) **41** (2004), no. 2, 185–204.

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES CA 90095-1555

*E-mail address*: tao@math.ucla.edu